

TARTU ÜLIKOOL

LOODUS- JA TEHNOLOOGIATEADUSKOND

MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT

BIOTEHNOLOOGIA ÕPPETOOL

Anna-Maria Himma

**Rasedusaegsete komplikatsioonide riskiga seonduvate geenide
koopiarvu varieeruvused**

Bakalaureusetöö

Juhendajad: Laura Kasak, MSc

Prof. Maris Laan

TARTU 2015

SISUKORD

SISUKORD	2
KASUTATUD LÜHENDID	4
SISSEJUHATUS	6
1. KIRJANDUSE ÜLEVAADE	8
1.1 Rasedusaegsed komplikatsioonid	8
1.2 Koopiaarvu varieeruvused.....	10
1.2.1 CNV-de tekkemehhanismid	12
1.2.2 CNV-de detekteerimine	13
1.2.2.1 Lookuspõhised meetodid.....	14
1.2.2.2 Ülegenoomsed meetodid	16
1.2.3 CNV-de seos haigustega	18
2. TÖÖ EESMÄRGID	20
3. TÖÖ ÜLESEHITUS	21
4. MATERJAL JA METOODIKA	22
4.1 Rasedusaegsete komplikatsioonide valim.....	22
4.2 DNA eraldamine platsentast.....	23
4.3 DNA eraldamine verest	24
4.4 Eksperimentaalne validatsioon <i>TaqMan</i> qPCR meetodil.....	25

4.5 Eksperimentaalne validatsioon <i>EvaGreen qPCR</i> meetodil	26
4.6 Andmete analüüs	28
5. TULEMUSED	29
5.1 Tuvastatud CNV-de valik eksperimentaalseks validatsiooniks ja piirkondade iseloomustus	29
5.2 Patsientide detailsem CNV-de profiil.....	32
5.3 Kiibitulemuste eksperimentaalne validatsioon <i>TaqMan qPCR</i> meetodiga	34
5.4 Replikatsiooniuuring laiendatud valimis	37
6. ARUTELU	41
6.1 Metoodika usaldusväärsus.....	41
6.2 Uuritud CNV-d ja nende seos haigustega	42
KOKKUVÕTE	45
SUMMARY	46
KIRJANDUSE LOETELU.....	48
KASUTATUD VEEBIAADRESSID	56
Lihthitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	57

KASUTATUD LÜHENDID

aCGH – võrdlev genoomne hübridisatsioon

CN – koopiaarv

CNV – koopiaarvu varieeruvus

Del – deletsioon

dNTP – desoksüribonukleotiidid

Dupl – duplikatsioon

EDTA – etüleendiamiintetraatseethape

FoSTeS – replikatsioonikahvli peatumine ja matriitsi ümbervahetus

GD – gestatsioonidiabeet

KMI – kehamassiindeks

LGA – suur gestatsiooniaja kohta

MAPH – multipleksne amplifitseeritava proovi hübridisatsioon

MLPA – multipleksne ligatsioon vahendatud proovi amplifikatsioon

MMBIR – mikrohomoloogia vahendatud katkemisest põhjustatud replikatsioon

NAHR – mittealleelne homoloogne rekombinatsioon

NHEJ – mittehomoogne otsade liitmine

PE – preeklampsia

PCR – polümeraasi ahelreaktsioon

*q*PCR – kvantitatiivne polümeraasi ahelreaktsioon

REPROMETA – REPROgrammed fetal and/or maternal METAbolism

SGA – väike gestatsiooniaja kohta

SNP – ühenukleotiidne polümorfism

SISSEJUHATUS

Sünnitajate keskmine vanus Eestis on 1992.-2013. aastatel suurenenud 25.5 aastalt 29.7 aastani. Esmasünnitajate keskmine vanus on tõusnud 22.7 aastalt 26.9 aastani, korduvsünnitajatel 28.3 aastalt 31.8 aastani. Koos sünnitajate vanuse kasvuga on tõusnud ka isade vanus (Allvee ja Karro, 2014). Andmed näitavad, et üha enam lükatakse lapse sünd edasi hilisemasse eaperioodi, mida põhjendatakse näiteks finantsprobleemide, sooviga olla rahaliselt kindlustatud, noore ea, majutusprobleemide, partneri puudumise või tuleviku ebakindlusega (Kutsar, 2012). Kõrgem rasedusaegne vanus tõstab aga rasedusaegsete komplikatsioonide riski (Luke ja Brown, 2007).

Raseduse teisel poolel võivad esineda mitmed tüsistused, nagu makrosoomia, hüpotroofia, gestatsioonidiabeet ja preeklampsia. Makrosoomia puhul esineb vastsündinutel gestatsiooniaja kohta liiga suur sünnikaal ning seda mõjutavad mitmed faktorid, osad geneetilised ja teised keskkonnast mõjutatud (Alberico *et al.*, 2014). Lisaks esineb seda tihti koos gestatsioonidiabeediga (Vora ja Bianchi, 2009), mille puhul ei talu naine suhkruid (Ferrara *et al.*, 2014). Haiguse vältimiseks tuleks naistel hoida tervislikku kehakaalu, ennast treenida ning hoiduda suitsetamisest (Zhang *et al.*, 2014), sest kõrge glükoositase kahjustab loodet nii ehituslikult kui funktsionaalselt (Zhang *et al.*, 2014).

Omavahel on seotud ka preeklampsia ning hüpotroofia (Hinkle *et al.*, 2014). Vastupidiselt makrosoomiale esineb hüpotroofia lastel gestatsiooniaja kohta väike sünnikaal ning riskideks on naise vähene kaal enne rasedust ja kaalukasv raseduse ajal. Esmasel rasedusel on risk hüpotroofiaks 9.5% (Hinkle *et al.*, 2014). Preeklampsiat põdevatel naistel esineb rasedusaegne kõrge vererõhk ning valgu kontsentratsioon uriinis (Zhao *et al.*, 2012). Hiliste rasedusaegsete komplikatsioonide kohta on teostatud mõned üksikud koopiaarvu varieeruvuste uuringud.

Koopiaarvu varieeruvused on vähemalt 50 aluspaari pikkused deletsioonid, duplikatsioonid, insertioonid, inversioonid ja translokatsioonid (Weischenfeldt *et al.*, 2013). Need on võrdlemisi levinud, moodustades 4.8-9.5% genoomist (Zarrei *et al.*, 2015). Enamus koopiaarvu varieeruvustest on päritud vanematelt, kuid nad võivad tekkida ka *de novo* (Strassberg *et al.*, 2011). Tekkemehhanisme on mitmeid, nii rekombinatsiooni-, replikatsiooni- kui ka reparatsioonipõhiseid (Weischenfeldt *et al.*, 2013). Juhul kui koopiaarvu varieeruvused katavad

üle miljoni aluspaari genoomis ning sisaldavad geenirikkaid alasid, võivad nad põhjustada erinevaid haigusi (Savage *et al.*, 2011).

Käesolevas bakalaureusetöös kirjeldatakse hiliskäigus esinevate rasedusaegsete komplikatsioonide fenotüüpi, riskifaktoreid. Lisaks iseloomustatakse koopiaarvu varieeruvusi, nende tekkemehhanisme, seost haigustega ning meetodeid nende analüüsiks genoomis.

Praktilise töö eesmärgiks oli varasemalt kiibikatsest välja valitud koopiaarvu varieeruvustega seotud piirkondade uurimine ja kirjeldamine määramaks, kas need on seotud erinevate rasedusaegsete komplikatsioonidega.

Märksõnad: rasedusaegsed komplikatsioonid, koopiaarvu varieeruvus (CNV), kvantitatiivne PCR (*qPCR*).

1. KIRJANDUSE ÜLEVAADE

1.1. Rasedusaegsed komplikatsioonid

Enamus rasedustest ja sünnitustest lõppevad heade tervisenäitudega nii emale kui lapsele, samas tüsistustega kulgenud rasedused võivad põhjustada ema ja/või lapse haigestumist või surma. Platsenta remodelleerimine lõpeb 16-18 gestatsiooninädalal ning selle ajal moodustatakse emaka ja platsenta vahelisest tsirkulatsioonist madala takistuse ja suure mahutavusega süsteem. Halva platsentatsiooni puhul võib tekkida hulgaliselt raseduskomplikatsioone (Khong et al., 2015). Raseduse teisel poolel võivad esineda mitmed tüsistused, nagu näiteks ema (preeklampsia ja gestatsioonidiabeet) ning last (suur või väike sünnikaal) mõjutavad komplikatsioonid.

Gestatsioonidiabeet (GD) on rasedusaegne suhkrutalumatus (Ferrara *et al.*, 2014). 2008.-2013. aastatel on Eesti rasedatel naistel esinenud gestatsioonitõbe 1.1-4.4% (Allvee ja Karro, 2014). Nendel naistel on seitse korda kõrgem risk haigestuda diabeeti ka pärast sünnitust. Tervisliku elustiili järgimine võib seda riski vähendada (Ferrara *et al.*, 2014). Tuleks hoida tervislikku kehakaalu ja toitumist, regulaarselt liikuda ning mitte suitsetada (Zhang, C *et al.*, 2014). GD tõstab nii ema kui lapse puhul komplikatsioonide riski raseduse ajal, sünnitusel ja edaspidi (Buckley *et al.*, 2012). Kõrge glükoositase mõjutab loote ehitust, funktsiooni ja arengut. Võivad esineda sünnidefektid, varases täiskasvanueas glükoositalumatus ning ülekaalulisus (Zhang, C *et al.*, 2014). Varajane haiguse detektsioon ja ravi vähendavad selle mõju (Buckley *et al.*, 2012). GD puhul ekspresseeruvad normaalsest erinevalt enim geenid, mis on seotud vastusega orgaanilisele materjalile, toitainete tasemele, hormoonidele ja endogeensele stiimulile; keemilise, glükoosi, karbohüdraatide tasakaalu hoidmisega; hormoonitaseme reguleerimisega ning heksooside metabolismiga (Zhang, Q *et al.*, 2015).

Gestatsiooniaja kohta suurt sünnikaalu (LGA, *Large for Gestational Age*) ehk makrosoomiat põhjustavad nii geneetilised kui keskkonnafaktorid (Alberico *et al.*, 2014). Põhiline LGA lapse sünni põhjus on gestatsioonidiabeet (Vora ja Bianchi, 2009). Vastsündinu makrosoomiat on seostatud veel ema ülekaalulisuse, liigse rasedusaegse kaalutõusu, diabeedi, ema pikkuse ja gestatsiooni vanusega sünnituse hetkel. Kui ema on ülekaaluline ja rasedusaegne kaalukasv soovituslikust madalam (7.0-11.5 kg), väheneb risk LGA lapse sünniks, kuid tõuseb risk väikese sünnikaaluga lapse sünniks (Alberico *et al.*, 2014). Makrosoomia puhul on suurenenud risk komplikatsioonideks sünnituse ajal, näiteks õlgade düstokia, või pärast sünnitust, näiteks

respiratoorse distressi sündroom, hüperbilirubineemia ja metaboolsed sündroomid (Rossi *et al.*, 2013). Võrreldes normaalsete platsentadega, on LGA puhul erinevalt ekspresseeritud geenid seotud lipiidide metabolismiga, väikeste molekulide biokeemiaga, organismi arenguga, rakulise arengu, kasvu ja jagunemisega ning kasvajate morfoloogiaga (Ahlsson *et al.*, 2015).

Gestatsiooniaja kohta väikese sünnikaalu (SGA, *Small for Gestational Age*) ehk hüpootroofia puhul on sünnikaal alla 10 protsentiili (Hodgetts *et al.*, 2014). SGA lapse sündimise riskifaktorid on ema madal rasedusaegne kaal, vähene kaalu kasv raseduse ajal ja rasedusaegne suitsetamine. Eri populatsioonides on välja toodud ka ema vanus, rasedustevaheline intervall, hüpertensioon ja sotsiaalmajanduslik staatus. Esmasünnitajal on SGA-ga lapse saamise tõenäosus 9.5%. Sellisel juhul esineb haigus teise raseduse puhul 80% tõenäosusega uuesti. Väikese sünnikaaluga lapse sünniks on suurem tõenäosus preeklampsiaga naistel (Hinkle *et al.*, 2014). SGA puhul on suurem risk loote haigestumiseks ja suremuseks ning hilisemas elus tüüp II diabeediks, hüpertensiooniks, rasvumiseks, südame-veresoonkonna haigusteks ja vaimse tervise häireteks (Hodgetts *et al.*, 2014). SGA teket seostatakse geenidega, mis vahendavad raku kasvu, kasvuhormooni sekretsiooni, insuliini tundlikkust, puberteedi hilinemist, rasvumist ja tüüp II diabeedi teket (Saenger ja Reiter, 2012).

Preeklampsia (PE) on komplikatsioon, mis 2009.-2013. aastatel on esinenud 2.1-2.7%-l Eesti rasedatest naistest (Allvee ja Karro, 2014). Seda iseloomustavad kõrge vererõhk (süstoolne üle 140 mmHg, diastoolne üle 90 mmHg) ning proteiinuuria (uriinis valgu kontsentratsioon üle 0.3 g ööpäevas). Varasemalt oli raske PE kriteeriumiteks Eestis vererõhk üle 160/110 mmHg ja/või proteiinuuria üle 5 g ööpäevas. Praegu on nendeks vererõhk üle 160/110 mmHg või mõni järgnevatest sümptomitest: trombotsüütide vähesus, maksafunktsiooni häired, tugev ravimitele mittealluv valu ülakõhus, neerupuudulikkus, kopsuturse, peaaaju- või nägemishäired (Woelkers *et al.*, 2015).

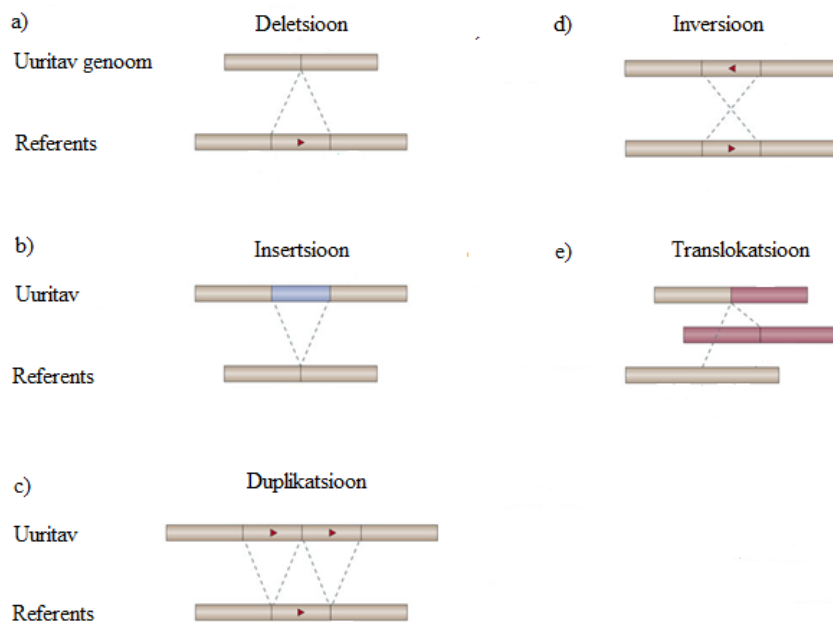
PE on põhiline rasedate haigestumise ja suremuse põhjus maailmas. Kuigi haiguse põhjused ilmuvad varem, tekivad kliinilised märgid alles raseduse teises pooles (Lambert *et al.*, 2014). On palju kandidaatgeene, mis võiksid olla seotud PE-ga. Need geenid on seotud nõrga immuunsüsteemi, platsenta isheemia ja suurenenud oksüdatiivse stressi taseme moodustumisega. Kandidaatgeenide selektsioon on limiteeritud, kuna ei mõisteta piisavalt PE bioloogilisi protsesse (Zhao *et al.*, 2012).

Preeklampsia jagatakse kaheks. Varase preeklampsia puhul toimub raseduse esimesel poolel ebatõhus platsentatsioon. Seeläbi on platsenta paljastatud hüdrostaatilisele ja oksüdatiivsele stressile, mis kahjustab rakke nii seest kui väljast. Haiguse kliinilisi märke kohe ei esine, samas detekteeritakse patoloogia enne 34-ndat rasedusnädalat. Hilisem preeklampsia on põhjustatud faktoritest, mis vabastatakse platsentast oksüdatiivse stressi mõjul. Normaalse raseduse puhul vabastatakse põletikureaktsioon samuti, eriti kolmanda trimestri lõpus, aga preeklampsia puhul on reaktsioon intensiivsem (Redman, 2011).

Kõigi hiliste rasedusaegsete komplikatsioonide kohta on kirjanduses ühenukleotiidsed polümorfismi (SNP, *single nucleotide polymorphism*) uuringuid (Harmon *et al.*, 2014; Chawla *et al.*, 2014; Beltcheva *et al.*, 2014), kuid koopiaarvu varieeruvuste (CNV, *copy number variation*) kohta vähe.

1.2 Koopiaarvu varieeruvused

CNV-de mõjul tekib DNA-d kas juurde või väheneb. Neid defineeritakse kui genoomseid segmente, mis on vähemalt 50 aluspaari pikad ja mille koopiaarv erineb, kui võrrelda kahte või enam genoomi (Zarrei *et al.*, 2015). Koopiaarvu varieeruvused võib lugeda insertioonideks, deletsioonideks, duplikatsioonideks, inversioonideks ja translokatsioonideks (Weischenfeldt *et al.*, 2013; joonis 1). Leidub hulgaliselt CNV-sid, mis esialgselt klassifitseeriti tähtsusetuteks, kuid nüüd arvatakse osalevat mõne haiguse vastuvõtlikkuses (Zarrei *et al.*, 2015).



Joonis 1. CNV-d on a) deletsioonid, b) insertioonid, c) duplikatsioonid, d) inversioonid ja e) translokatsioonid (Weischenfeldt *et al.*, 2013 järgi).

Autosoomides peaks igat segmenti olema kaks koopiat, üks kummagil kromosoomil. Deletsiooni puhul aga esineb vaid üks või 0 koopiat ning duplikatsioonil kolm või enam (Savage *et al.*, 2011). CNV-d moodustavad 4.8-9.5% inimese genoomist, kusjuures 3.6-7.5% nendest moodustavad deletsioonid (Zarrei *et al.*, 2015). Umbes 65-80% inimestest kannavad ühte üle 100 000 aluspaari suurust CNV-d, 5-10% vähemalt 500 000 aluspaari suurust ja 1% üle 10^6 aluspaari suurust CNV-d (Girirajan *et al.*, 2011). Umbes 41% CNV-dest kattuvad geenidega, mis näitab, et nad võivad mõjutada nende ekspressiooni. Enamik koopiarvu varieeruvustest päritakse vanematelt, osa tekib *de novo* (ei ole päritud kummaltki vanemalt) (Strassberg *et al.*, 2011). CNV-d võivad moodustuda DNA rekombinatsiooni, replikatsiooni ja reparatsiooni käigus (Weischenfeldt *et al.*, 2013).

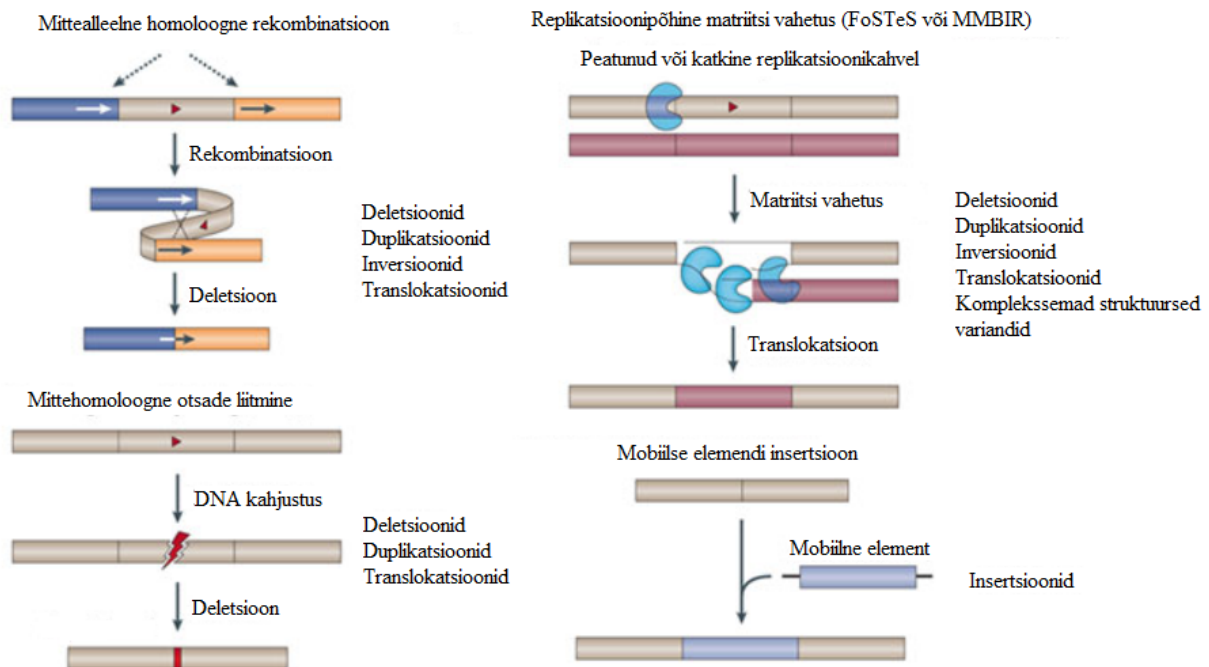
1.2.1 CNV-de tekkemehhanismid

Põhilised CNV-de tekkemehhanismid on (joonis 2):

1. Mittealleelne homoloogne rekombinatsioon (NAHR, *Nonallelic Homologous Recombination*) (Liu, P. *et al.*, 2012). NAHR toimub tavaliselt kahe sama päritolu madala koopiaarvuga korduste ja segmentaalsete duplikatsioonide vahel. Mehhanismi tulemusel võivad tekkida deletsioonid, duplikatsioonid, translokatsioonid ja inversioonid (Dittwald *et al.*, 2013).
2. Mittehomoloogne otsade liitmine (NHEJ, *Nonhomologous End-Joining*) (Liu, P. *et al.*, 2012). Kaheahelalisi katkeid põhjustavad V(D)J rekombinatsioon, ioniseeriv radiatsioon ja reaktiivsed hapnikuühendid. NHEJ parandab DNA kaheahelalisi katkeid, liites kromosoomi otsad kokku. Tavaliselt tekivad produktidesse deletsioonid (Waters *et al.*, 2014).
3. Replikatiivsed mehhanismid
 - a. Replikatsioonikahvli peatumine ja matriitsi ümbervahetus (FoSteS, *Fork Stalling and Template Switching*) (Liu, P. *et al.*, 2012). FoSteS meetodi puhul replikatsioonikahvel peatub, mahajääv ahel tuleb algkohast lahti ning ühendub mõne teise replikatsioonikahvliga. Uus ahel ei pea olema algse replikatsioonikahvliga kõrvuti, kuid siiski lähedal. Mehhanismi järgselt tekivad kas duplikatsioonid, deletsioonid, inversioonid või komplekssemad CNV-d (Zhang *et al.*, 2009).
 - b. Mikrohomoloogia vahendatud katkemisest põhjustatud replikatsioon (MMBIR, *Microhomology-Mediated Break-Induced Replication*) (Liu, P. *et al.*, 2012). MMBIR parandab katkenud kaheahelalisi otsi, kui replikatsioonikahvel peatub ja kokku variseb. Kasutatakse õdekromatiidi või homoloogse kromatiidi replikatsioonikahvli. Mehhanismi tõttu tekivad genoomi deletsioonid või duplikatsioonid (Sahoo *et al.*, 2015).
4. Pikad vahelduvad elemendid
 - a. Vahendatud retrotranspositsioon (Liu, P. *et al.*, 2012). L-1 retrotranspositsioon algab mRNA transkriptsioonist. mRNA viiakse tsütoplasmasse ja transleeritakse. Moodustub L1 ribonukleopartikkel, mis siseneb tuuma ja sünteesib cDNA koopia,

mis integreerub genoomi. Sama ensüümikompleksi abil saavad mobiliseeruda ka teised retrotransposonid (Carreira *et al.*, 2014).

- b. Mobiilse elemendi insertioon (Liu, P. *et al.*, 2012). Kui retrotransposonid paljunesid RNA vahendusel ja kasutasid kopeerimise ja kleepimise meetodit, siis transposonid relokeeruvad genoomses DNA-s ühest kohast teise (Carreira *et al.*, 2014).



Joonis 2. CNV-de viie tekkemehhanismi iseloomustus: mittealleelne homologne rekombinatsioon, mobiilse elemendi insertioon, replikatsioonikahvli peatumine ja matriitsi vahetus, mikrohomoloogia vahendatud katkemisest põhjustatud replikatsioon ja mittehomoloogne otsade liitmine. Paremal on toodud iga mehhanismiga tekkivad koopiarvu varieeruvuste tüübid (Weischenfeldt *et al.*, 2013 järgi).

1.2.2 CNV-de detekteerimine

CNV-sid võib tuvastada erinevate analüüsivormidega, näiteks võrdleva genoomse hübriidisatsiooni (Kallioniemi *et al.*, 1992), SNP genotüpiseerimisega (Korn *et al.*, 2008) ja polümeraas ahelreaktsioonil (PCR, polymerase chain reaction) (Kellogg *et al.*, 1990) põhinevate

meetoditega. Erinevad meetodid ja algoritmid annavad erinevaid tulemusi (Zhang, F. *et al.*, 2009) ja neil on nii positiivsed kui negatiivsed küljed (Vandeweyer ja Kooy, 2013). Antud bakalaureusetöös kasutati kvantitatiivset PCR-i (*qPCR*, *Quantitative Polymerase Chain Reaction*) ja eelnevalt oli läbi viidud ülegenoomne SNP mikrokiibi uuring, mistõttu kirjeldatakse neid meetodeid pikemalt.

1.2.2.1 Lookuspõhised meetodid

On olemas mitmeid PCR-il põhinevaid kvantitatiivse detektsiooni meetodeid. Kõige lihtsaim neist on *qPCR*, kus reaktsiooni produktid märgistatakse fluorestsentsvärvidega. Et vähendada reaktsiooniks vajaminevaid komponente, arendati välja ka teisi meetodeid: multipleksne amplifitseeritava proovi hübridisatsioon (MAPH, *Multiplex Amplifiable Probe Hybridisation*) (Armour *et al.*, 2000) ja multipleksne ligatsioon-vahendatud proovi amplifikatsioon (MLPA, *Multiplex Ligation-Dependent Probe Amplification*) (Schouten *et al.*, 2002). MAPH puhul kinnitatakse genoomne DNA nailonmembraanile, märgistatud proovid hübridiseeritakse sellele, amplifitseeritakse ning võrreldakse nende intensiivsust. Multipleksimine on võimalik, kui kasutada eri pikkustega proove, mis on sama praimeriga seondumissaidiga. Produktid eraldatakse pikkuse järgi, ning kuna kõik uuritavad sisaldavad sama praimeriga seondumissaiti, mõõdetakse intensiivsust lõpp-punktist. MLPA ei kasuta immobilisatsiooni. Selle meetodiga hübridiseeritakse kaks proovi DNA-le ning seejärel ligeeritakse, amplifitseeritakse ja märgistatakse. Eri pikkusega produktid eraldatakse üksteisest ja mõõdetakse lõpp-punkti intensiivsust (Vandeweyer ja Kooy, 2013).

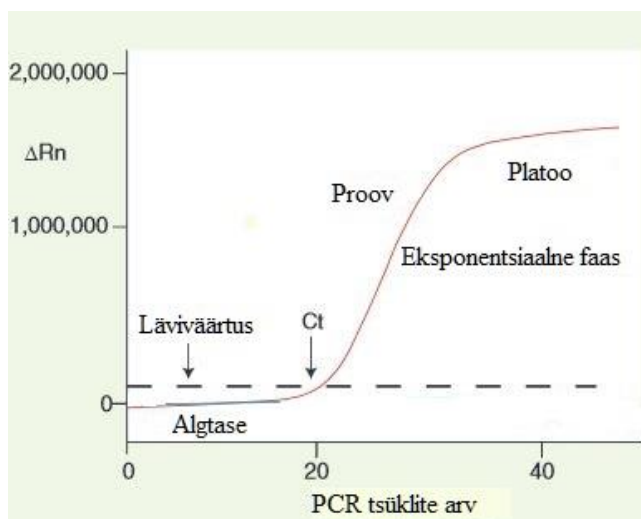
PCR-i reaktsioon moodustab DNA koopiaid eksponentsiaalselt. Kui tavalise PCR-i puhul lõpptulemused varieeruvad, siis kvantitatiivse PCR-i puhul mõõdetakseprodukte iga PCR-i tsükli ajal. *qPCR*-il on erinevaid rakendusvõimalusi: mRNA ekspressiooniuuring, splaissingu produkti ekspressiooni analüüs, aga ka DNA koopiaarvu mõõtmine genoomses ja viiruslikus DNA-s (Arya *et al.*, 2005).

Kõige enam kasutatakse *qPCR*-i puhul *TaqMan* proove (Mao *et al.*, 2007) ning termostabiilset ensüümi *Thermus aquaticus* (*Taq*) polümeraasi, millel on 5' – 3' eksonukleasne aktiivsus. Seda aktiivsust kasutatakse ära produkti detekteerimisel. *TaqMan* proov hübridiseerub sihtmärk DNA-

le, on märgistatud ning ei käitu praimerina. Sellel on algselt 5' otsas reportervärv ja 3' otsas vaigistaja. Vaigistaja lähedus vähendab reprotervärvilt poolt väljakiirutatavat fluorestsentsi. Kui proov kinnitub PCR-i produktile, kasutab *Taq* polümeraas eksonukleaasest aktiivsust, muutes selle väiksemateks fragmentideks. Reporter ja vaigistaja eemalduvad üksteisest ning fluorestsentssignaal tugevneb. PCR-i produkt on nüüd vaba ning selle järgi hakatakse praimerit pikendama (Arya *et al.*, 2005).

Lisaks on võimalik kasutada värve, mis seonduvad kaheaahelalise DNA-ga mittespetsiifiliselt. Fluorestsentssignaal tugevneb seetõttu polümerisatsiooni käigus ning väheneb denaturatsioonis. Mõõtmised tehakse iga elongatsiooni lõpus. Selle tehnika eelis on hind, kuna värvi saab kasutada iga praimeripaariga, samas pole meetod sihtmärk-DNA suhtes spetsiifiline (Arya *et al.*, 2005). Üheks sellise värvi näiteks on *EvaGreen* (Mao *et al.*, 2007).

Ükskõik millist värvi kasutatakse, detekteerib termotsükler fluorestsentsi emissiooni ning moodustab sellest amplifikatsioonikõvera (joonis 3). Algtasemeks on PCR-i tsüklid, kus fluorestsentssignaal tugevneb, aga on veel liiga nõrk, et masin suudaks seda detekteerida. ΔR_n väärtuse arvutab välja masin, kasutades võrrandit $R_n = R_{nf} - R_{nb}$, kus R_{nf} on produkti emissioon iga ajahetkes ja R_{nb} läviväärtuse emissioon. Signaal, mis detekteeritakse peale läviväärtust, loetakse õigeks ning seda võib kasutada C_t defineerimiseks. C_t on PCR-i tsükli arv, kus fluorestsents on tugevam kui minimaalne detektsioonitase. C_t -d määratakse eksponentsiaalses faasis, kuna selles faasis pole reaktsioonikomponendid limiteeritud (Arya *et al.*, 2005).



Joonis 3. *q*PCR-i amplifikatsioonikõver. Absoluutse kvantifikatsiooni puhul tahetakse teada uuritava proovi kogust. Selleks mõõdetakse C_t ning kasutades standardkõverat määratakse koopiaarv. Standardkõver moodustatakse lahjendusrea abil. Relatiivse kvantifikatsiooni puhul seda vaja pole. Tulemused arvutatakse välja kalibraatori suhtes valemiga $\Delta\Delta C_t = \Delta C_t(\text{uuritav}) - \Delta C_t(\text{kalibraator})$ (Arya et al., 2005).

1.2.2.2 Ülegenoomsed meetodid

Põhilised kasutusel olevad kogu genoomi hõlmavad meetodid on SNP ja võrdleva genoomse hübriidatsiooni (aCGH, *Array Comparative Genome Hybridisation*) mikrokiibid (Li ja Olivier, 2013). aCGH-ga kasutatakse erinevalt märgistatud uuritavat ja referents-DNA-d, mis hübriidiseeritakse koos samale mikrokiibi punktile (Vanderweyer ja Kooy, 2013). Mõõdetakse fluorestsentsi tasemed koos iga kromosomaalse regiooni pikkusega ning neid võrreldes detekteeritakse uuritavas DNA-s duplikatsioonid või deletsioonid. aCGH mikrokiibid kasutavad kas bakteri tehislikke kromosome, cDNA-d või pikki sünteetilisi oligonukleotiide sondeerimaks kindlaid DNA regioone. aCGH on ülegenoomsetest meetoditest kõrgeima tundlikkuse ja spetsiifilisusega, aga ruumiline resolutsioon ning läbilaskevõime on madalad (Li ja Olivier, 2013).

Iga inimene pärib kummaltki vanemalt ühe alleeli, nii et genotüüp SNP saidis on kas AA, BB või AB. Kuna SNP-d on laialt levinud üle kogu genoomi, saab nende abil uurida genoomi muutusi ja

ümberkorralduste kindlaid asukohti. Katses kasutatav fragmenteeritud ja märgistatud sihtmärk-DNA hübridiseeritakse kiibile. Iga oligonukleotiidi sisaldab SNP piirkonda ning nad on spetsiifilised mõlemale SNP alleelile, A ja B. Katses määratakse mõlema alleeli fluorestsentssignaal, mis annab infot nii SNP genotüübi (A, B, AB) kui ka koopiaarvu kohta (heterosügootsus või homosügootsus/hemisügootsus) (Iacobucci *et al.*, 2013).

SNP mikrokiipide kasutamise eelisteks on testi madal hind ning väike vajaminev DNA hulk (Winchester *et al.*, 2009). Kiibile on võimalik kinnitada üle kuue miljoni oligonukleotiidi, mis on seotud kas kindla haigusega või iseloomustavad tervet genoomi (Savage *et al.*, 2011). Kaks põhilist SNP mikrokiipide tootjat on Affymetrix, Inc (Santa Clara, CA, USA) ja Illumina (San Diego, CA, USA). Mõlemad kasutavad koopiaarvu määramiseks ära samu biokeemilisi printsiipe, kuid erinevat keemilist meetodit (Iacobucci *et al.*, 2013).

Illumina meetod põhineb platformil, kus kogu genoom amplifitseeritakse, fragmenteeritakse ja hübridiseeritakse kuulidele, mis sisaldab igale SNP lookuse alleelile spetsiifilisi oligonukleotide. Affymetrix'i tehnoloogiaga DNA lõhustatakse ja seejärel amplifitseeritakse, märgistatakse ja hübridiseeritakse mikrokiibile (Iacobucci *et al.*, 2013).

Fluorestsentssignaalide põhjal proovid analüüsitakse. Selleks on arendatud mitmeid algoritme. Uuritava proovi SNP genotüüpe tuleb võrrelda referentsi omadega. Andmeid kirjeldatakse \log_2 väärtustena, mis arvutatakse uuritava ja referents-DNA intensiivsuse suhtena. Kui $\log_2=0$, siis pole koopiaarvus mingeid muutuseid, ning kui \log_2 on 0-st väiksem või suurem, on vastavalt tegemist deletsiooni või duplikatsiooniga (Iacobucci *et al.*, 2013).

CNV-de detekteerimiseks genotüpiseerimisandmetest on olemas suur hulk erinevaid algoritme ja tarkvarasid. Nendeks on näiteks Birdsuite 1.5.5, CNAT, GADA, PennCNV ja QuantiSNP (Winchester *et al.*, 2009). Kidd *et al.* (2008) näitasid, et nendest kõige enam CNV-sid detekteeris GADA algoritm, kus SNP mikrokiibi analüüs viidi läbi Affymetrix 6.0 meetodiga.

1.2.3 CNV-de seos haigustega

CNV-d asuvad eelistatult geenidest väljaspool, ülikonserveerunud piirkondades ning haigustega seotud geenides esineb pigem duplikatsioone kui deletsioone (Zhang, F *et al.*, 2009). Duplikatsioonide ja deletsioonide suhe inimeste verest eraldatud DNA-s on 2.34, mis viitab positiivsele selektsioonile primaatide genoomi evolutsioonis (Dumas *et al.*, 2007). CNV-d põhjustavad nii normaalset fenotüübi varieeruvust kui ka haigustele vastuvõtlikkust (Zhang, F *et al.*, 2009).

CNV-d võivad olla seotud haigustega erinevate mehhanismide mõjul: näiteks geeni koguse, liitumise või positsiooniefektide mõjul (Zhang, F *et al.*, 2009). Kõik koopiarvu varieeruvused ei mõjuta geeniekspressiooni, mis tähendab, et on olemas kas doosi kompensatsioonimehhanismid või ei hõlma CNV kõiki regulaatorelemente. Ekspressioon võib tõusta, kui deleteeritakse näiteks transkriptsiooni repressor (Liu, G *et al.*, 2014). Põhilised haiguseid tekitavad CNV-d on pikemad kui miljon aluspaari ja sisaldavad endas geenirikkaid alasid (Savage *et al.*, 2011).

Esineb kahte tüüpi haiguseid tekitavaid koopiarvu varieeruvusi. Esimesed neist esinevad enam kui ühes protsendis populatsioonist ning koopiarv varieerub nullist 30-ni. Nende puhul on oluline mõista koopia sisu või järjestuse struktuuri. Teised on haruldased (alla 1%), esinedes vähemate koopiatena ja olles tugeva loodusliku valiku all. Nad tekivad sagedaselt *de novo* ning püsivad vaid paar põlvkonda (Girirajan *et al.*, 2011). CNV-dega on seostatud peamiselt selliseid haiguseid nagu autism, skisofreenia, Crohni tõbi ja raske varane ülekaalulisus (Liu, G. *et al.*, 2014).

Rasedusaegsete komplikatsioonidega seotud CNV uuringute tulemusi kirjeldavaid artikleid on ilmunud vaid preeklampsia ning hüpotroofia kohta. Zhao *et al.*, 2012 teostasid SNP mikrokiibi analüüsi 117 preeklampsiaga naisel ja 116 kontrollil. Leiti kolm harvaesinevat CNV-d, mis olid rikastatud PE grupis. Neist kõige huvitavam oli 15 000 aluspaari pikkune deletsioon *PSG11* (rasedusspetsiifiline glükoproteiin) geenis. Rasedusspetsiifilisi glükoproteiine toodetakse raseduse ajal platsentas ja need moodustavad ühe immunoglobuliinide perekonna.

Canton *et al.*, 2014 viisid läbi kiibipõhise võrdleva genoomihübridisatsiooni katse, leidmaks 51-l SGA-ga sündinud lapsel CNV-sid. Leiti 18 CNV-d, mida tervetel indiviididel pole täheldatud.

Neist üheksa olid patogeensed ning kolm kattusid teadaolevate deletsioonidega. Need alad on geenirikkad, mõjutavad mitut kandidaatgeeni või on seotud kasvuregulatsiooniga.

Bustamante *et al.*, 2012 uuringus valiti kirjanduse põhjal 10 CNV-d ning genotüpiseeriti nende suhtes 884 lapse ja 913 ema DNA-d. Leiti, et *GSTT2B* geeni insertioon laste DNA-s tõstis SGA riski.

2. TÖÖ EESMÄRGID

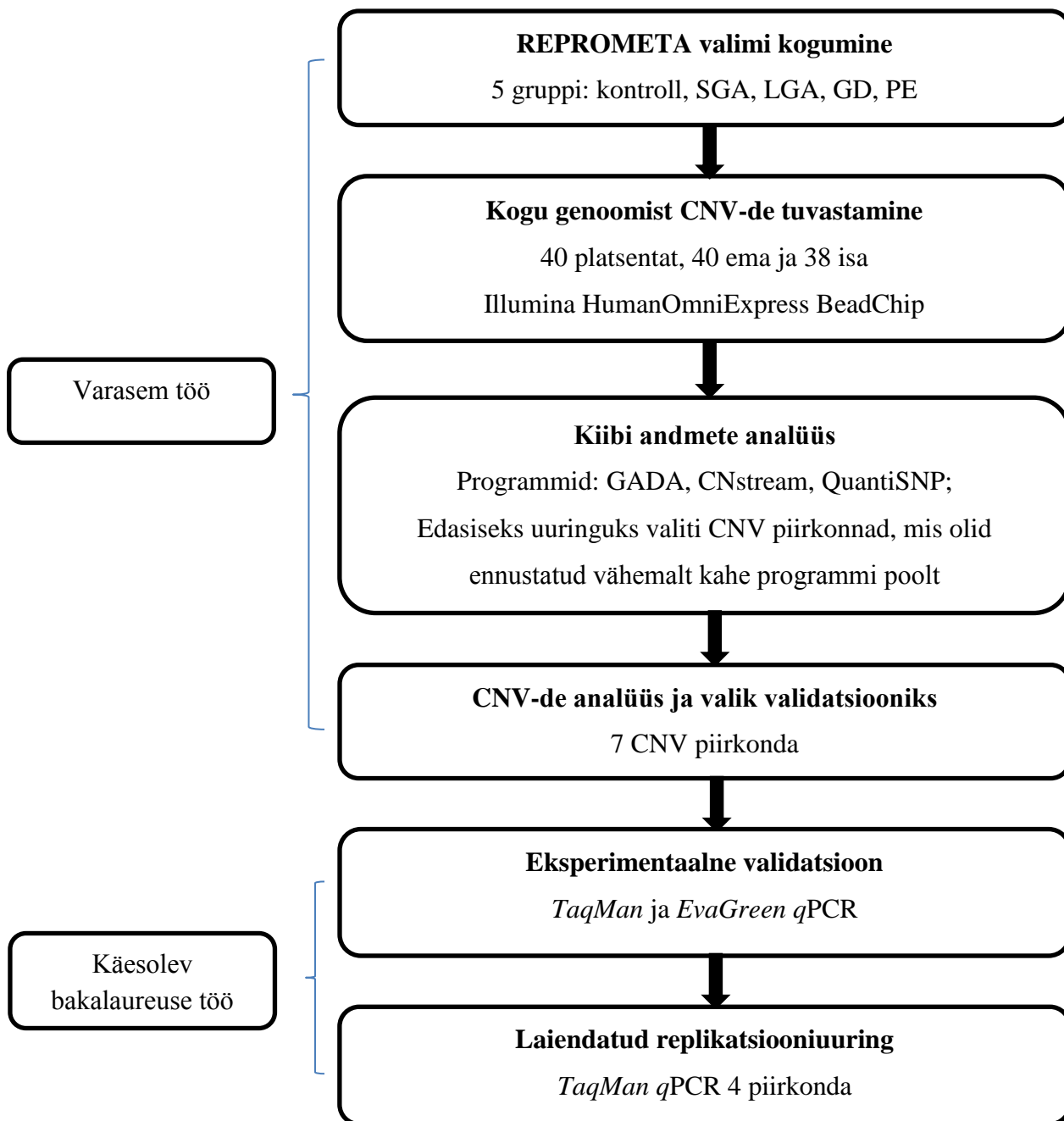
Käesolevale bakalaureusetööle eelnevalt teostati ülegenoomne SNP kiibiuuring trio materjalis: platsenta ning ema ja isa verest eraldatud DNA. Töö eesmärgiks oli tuvastada ja kirjeldada CNV piirkondi, mis võiksid olla seotud raseduse teises pooles esinevate raseduskomplikatsioonidega. Nendeks komplikatsioonideks olid gestatsioonidiabeet, preeklampsia, gestatsiooni aja kohta liiga suur ja väike sünnikaal.

Antud bakalaureusetöö eesmärgiks oli uurida eelnevalt välja valitud CNV piirkondade seost rasedusaegsete komplikatsioonidega. Täpsemad eesmärgid olid:

- a) varasemalt teostatud kiibiuuringu tulemuste kinnitamine *qPCR*-i meetodiga;
- b) replikatsiooniuringu läbiviimine, kinnitamaks koopiaarvu varieeruvuste seost rasedusaegsete komplikatsioonidega.

3. TÖÖ ÜLESEHITUS

Joonis 4 kirjeldab antud bakalaureusetöö raames teostatud katseid ja sellele eelnenud tööd.



Joonis 4. Ülevaade käesoleva bakalaureuse töö raames tehtud ning sellele eelnenud tööst. REPROMETA valimit kogusid Kristiina Rull, Pille Vaas ja Pille Teesalu. SNP mikrokiibi andmete analüüsi viis läbi Laura Kasak, piirkondade analüüsi ja valiku edasiseks uurimiseks tegid Laura Kasak ja prof. Maris Laan.

4. MATERJAL ja METOODIKA

4.1. Rasedusaegsete komplikatsioonide valim

Antud uuringus kasutati REPROMETA (REPROgrammed fetal and/or maternal METAbolism) valimit. REPROMETA valimit on kogutud alates 2006. aasta märtsist seoses projektiga „Platsenta hormoonide sünteesi määravate geenide ekspressiooni ja varieeruvuse seos raseduse kulu ja üsasisese arenguga“ (Männik *et al.*, 2010; Uusküla *et al.*, 2012). Projekti läbiviimiseks on saadud luba Tartu Ülikooli inimuuringu eetika komiteelt (loa numbrid 146/18, 27.02.2006; 150/33, 18.06.2006; 158/80, 26.03.2007). Valimi kogumist rahastati prof. Maris Laane grantidest HHMI #55005617, ETF 7471, WT 070191/Z/03/Z ja HTM 0182721s06. Kõik uuringus osalejad on allkirjastanud informeeritud nõusoleku lehe. Sünnitusjärgselt võeti platsentakoe ja nabaväädi vereproovid ning ema ja isa vereproovid.

REPROMETA uuringu pered jaotati gruppidesse vastsündinute sünnikaalu ning ema rasedusaegsete komplikatsioonide (mitte)esinemise alusel. Käesolevas töös kasutati kontrollgruppi ning lapsega seonduvaid (makrosoomia, hüpotroofia) ja emapoolseid (gestatsioonidiabeet, preeklampsia) rasedustüsistusi. Kõikidest gruppidest kaasati 8 platsenta DNA-d ning vanemate verest eraldatud DNA-d (tabel 1). Replikatsioonikatses laiendati valimit sõltuvalt validatsioonikatse tulemustest.

Kontrollgrupp koosneb normaalse raseduskuluga rasedustest, kus vastsündinu kaal on normaalne – õigeaegselt sündinud laps, kelle sünnikaal on 10.-90. protsentiili vahel. Madala sünnikaaluga lastel oli sünnikaal gestatsiooniaja ja sünnikaalu kohta alla 10. protsentiili, suure sünnikaaluga vastsündinutel oli sünnikaal gestatsiooniaja ja sünnikaalu kohta üle 90. protsentiili. Protsentiilid arvutati Eesti Meditsiinilise Sünniregistri alusel.

Preeklampsia patsiendid valiti raske preeklampsiaga, kellel oli diagnoositud hüpertensioon - süstoolne vererõhk ≥ 160 mmHg, diastoolne vererõhk ≥ 110 mmHg ja/või proteinuuria ≥ 5 g ööpäevas. Preeklampsiat diagnoosisid patsientidel Kristiina Rull ja Pille Vaas. Gestatsioonidiabeet diagnoositi oraalse glükoosi tolerantsus testiga 24-28 gestatsiooninädalal, manustades patsiendile 75 g glükoosi. Veenivere glükoositaseme paastuplasmas >5.1 mmol/l ja/või 1h ja 2h järel > 10 mmol/l ja >8.5 mmol/l.

Tabel 1. Kiibikatses kasutatud valimi iseloomustus.

	Kontroll	SGA	LGA	PE	GD
Emad					
Vanus (aastad)	29.3 ± 7.85	24.5 ± 3.51	30.8 ± 3.51	27.3 ± 7.39	30.2 ± 5.15
KMI enne rasedust (kg/m ²)	23.9 ± 3.77	20.6 ± 2.92	24.8 ± 4.61	26.3 ± 5.00	26.3 ± 7.85
Kaaluiive raseduse ajal (kg)	17.2 ± 3.52	13.3 ± 3.47	21.7 ± 7.25	10.9 ± 3.36	14.8 ± 5.76
Gestatsiooni- vanus (päevades)	278.6 ± 11.54	272.0 ± 9.46	281.6 ± 4.41	266.1 ± 3.94	275.5 ± 5.98
Isad					
Vanus (aastad)	31.8 ± 5.99	27.1 ± 5.64	35.8 ± 8.28	32.8 ± 8.94	33.5 ± 7.05
KMI (kg/m ²)	25.1 ± 3.82	22.7 ± 2.04	29.9 ± 5.64	28.8 ± 6.34	27.8 ± 4.34
Laps					
Sünnikaal (g)	3702.8 ± 392.23	2442.1 ± 234.84	4726.2 ± 207.86	2793.8 ± 487.89	4268.5 ± 238.34
Sünnipikkus (cm)	51.3 ± 1.89	46.3 ± 1.04	53.4 ± 1.18	47.6 ± 1.51	52.4 ± 1.30
Platsenta kaal (g)	571.3 ± 115.81	397.5 ± 89.88	816.3 ± 115.97	476.9 ± 119.01	680.7 ± 187.80

Andmed on näidatud vormis keskväärtus ± standardhälve. KMI: kehamassiindeks; SGA: gestatsiooniaja kohta väike sünnikaal; LGA: gestatsiooniaja kohta suur sünnikaal; PE: preeklampsia; GD: gestatsioonidiabeet.

4.2 DNA eraldamine platsentast

Platsentast DNA eraldamiseks eemaldati platsentast tükk kaaluga 25 mg ja lõigati see väiksemateks tükkideks. Proov asetati mikrotsentrifuugi tuubi, kuhu lisati 180 µl puhver T1-te ja 25 µl Proteinaas K lahust. Segu segati vorteksiga, nii et kude oli täielikult kaetud lüüsilahusega. Seejärel hoiti proovi 56°C juures, kuni oli toimunud täielik lüüs (1-3 tundi või üleöö). Selle aja

jooksul kasutati kas aeg-ajalt vorteksit või loksutavat inkubaatorit. Seejärel proovi segati vorteksil, lisati 200 µl puhver B3-e ja segati vorteksil jõuliselt. Proovi inkubeeriti 70°C juures 10 minutit ja segati vorteksil. Lisati 210 µl (96-100%) etanooli ja segati jõuliselt vorteksil.

Iga proovi kohta asetati üks NucleoSpin® Tissue Column kogumistuubi ja valati proov kolonni. Tsentrifugiti 1 minut 11 000 x g. Läbi voolanud vedelik visati ära ja kolonn pandi tagasi kogumistuubi. Seejärel membraani pesti. Esimesel pesul lisati 500 µl puhver BW-d, tsentrifuugiti 1 min 11 000 x g-l ja eemaldati eluaat. Teisel pesul lisati 600 µl puhver B5-e, tsentrifuugiti 1 minut 11 000 x g ja eemaldati eluaat. Kolonni tsentrifuugiti 1 min 11 000 x g, selles etapis eemaldus jääk-etanool. Kolonn asetati 1.5 mL suurusesse tuubi ja lisati 100 µl eelsoojendatud (70°C) puhvrit BE. Proovi inkubeeriti toatemperatuuril 1 minut ja 1 min tsentrifuugiti 11 000 x g.

4.3 DNA eraldamine verest

DNA eraldamiseks EDTA-verest lisati jääl 1 ml vere kohta 4 ml puhvrit A [0,32 M sahharoos, 1 mM Tris-HCl (pH 7,5), 5 mM MgCl₂, 1% Triton X-100, ddH₂O]. Tuubi keerati korralikult, et toimuks lüüs. Proovi tsentrifuugiti 30 min 3000 rpm +4°C juures. Seejärel kallati ära supernatant ning vajadusel korraliti sama tegevust. Pesemiseks lisati 2 ml Puhver A-d ning tsentrifuugiti 20 min 3000 rpm +4°C juures. Eemaldati supernatant, sademele lisati 0,5 ml puhvrit B [400 mM NaCl, 2 mM EDTA (pH 8), 10 mM Tris-HCl (pH 7,5), ddH₂O] ja toksiti sade lahusesse. Tilkhaaval lisati 100 µl puhvrit C [20% SDS, 20 mg/ml proteinaas K, ddH₂O] ja inkubeeriti üleöö 37°C.

Järgmisel päeval lisati inkubeeritud lahusele 270 µl küllastatud NaCl, mis seob valgud, ja raputati jõuliselt 25 sekundit. Proovi tsentrifuugiti 30 minutit 4000 rpm toatemperatuuril. Supernatant viidi teise tuubi ja lisati kahekordne kogus 96% etanooli. Proovi segati korralikult ilma vorteksita. Pasteuri pipeti otsaga võeti sealt DNA ning viidi 300-500 µl TE-sse (pH 7.5) [10 mM Tris-HCl (pH 7.5), 1 mM EDTA). Proovi hoiti 4-7 päeva +4°C juures loksutil.

4.4 Eksperimentaalne validatsioon *TaqMan qPCR* meetodil

Esmane eksperimentaalne validatsioon viidi läbi samade DNA proovidega, mille põhjal oli eelnevalt teostatud CNV-de tuvastamine kogu genoomist kiibipõhiselt (tabel 1, joonis 2). Validatsiooniks kasutati *TaqMan qPCR* meetodikat ehk CNV piirkondades asuvaid kommertsiaalselt kättesaadavaid valmis proove (*pre-designed copy number assay*, Applied Biosystems, USA). Validatsioonile eelnes optimeerimine, kus määrati igale *assay*-le sobivaim kogus ja parimad parameetrid, millega tulemusi analüüsida.

TaqMan qPCR viidi läbi bipleks reaktsioonidena (paralleelselt uuritav piirkond ja referentsgeen *RNaseP*) suhtelise kvantifitseerimise (*relative quantification*) meetodil. Reaktsioonisegu lõppmahus 10 µl sisaldas 2 µl 5 x HOT FIREPol® Probe *qPCR* Mix Plus'i (ROX) (Solis BioDyne, Eesti), mis sisaldas HOT FIREPol® DNA polümeraasi, 5 x Probe *qPCR* puhvrit, 15 mM MgCl₂, dNTP-sid, ROX värvainet, TaqMan® Copy Number Assay'd (Applied Biosystems, USA), TaqMan® Copy Number Reference Assay *RNase P*'d (Applied Biosystems, USA) (tingimused toodud tabelis 2), 1 µl genoomset DNA-d (10 ng/µl) ning ddH₂O-d lõppmahuni. Kõik reaktsioonid viidi läbi kolmes korduses kahe katsena. Kõikidele plaatidele lisati kvaliteedi kontrolliks üks referentsindiviid, kellel oli teadaolevalt 2 koopiat ja plaadisiseks normaliseerimiseks kasutatav viljakate kontrollide DNA-de segu. Kontrollide DNA segu valmistati 50-st juhuslikult valitud kontrollindiviidi DNA-st.

Tabel 2. Validatsiooniks valitud piirkondade asukohad kromosoomil, hõlmatud geeni nimed, kasutatud TaqMan® Copy Number Assay kood ning kasutatud referentsi *RNaseP* ja *TaqMan* proovi kogused 1 µl proovi kohta.

CNV asukoht	Geen	TaqMan proov	Referentsi kogus (µl)	<i>TaqMan</i> proovi kogus (µl)
3p21.1	<i>SFMBT1</i>	Hs03451854_cn	0.5	0.5
5q35	<i>NSD1</i>	Hs00046211_cn	0.5	0.5
7q34	<i>MGAM</i>	Hs03628107_cn	0.5	0.5
1q2	<i>RABGAP1L</i>	Hs04197428_cn	1	0.5
10p13	<i>CCDC3</i>	Hs03285499_cn	0.5	0.5
8q23	<i>SLC26A7</i>	Hs06218115_cn	0.5	0.5
11p13	<i>CAT</i>	Hs00351750_cn	0.5	0.5

Amplifikatsioonireaktsioonid viidi läbi ABI Prism 7900HT Sequence Detection System (Applied Biosystems, USA) masinaga järgmistel tingimustel:

Algne denaturatsioon	95°C – 15 min	
Denaturatsioon	95°C – 15 s	} 40 tsüklit
Praimerite seondumine ja ekstentsioon	60°C – 60 s	

4.5 Eksperimentaalne validatsioon *EvaGreen* qPCR meetodil

Validatsioon *EvaGreen* qPCR meetodiga viidi läbi samade DNA proovidega, mida kasutati *TaqMan* qPCR puhul. Katse teostati *TaqMan* qPCR meetodil läbi viidud validatsiooni tulemuste kinnitamiseks *CAT* piirkonnas. Selleks disainiti ennustatud CNV piirkondade praimerid (tabel 3). Validatsioonile eelnes optimeerimine, kus määrati praimerite paarile sobivaim kogus ja parimad parameetrid, millega tulemusi analüüsida.

EvaGreen qPCR viidi läbi unipleks reaktsioonidena (ühel plaadil eraldi uuritav piirkond ja referentsgeen *ALB* (albumiin)) absoluutse kvantifitseerimise (*absolute quantification*) meetodil. Reaktsioonisegu lõppmahus 10 µl sisaldas 2 µl 5 x HOT FIREPol® *EvaGreen*® qPCR Mix Plus'i (ROX) (Solis BioDyne, Eesti), mis sisaldas HOT FIREPol® DNA polümeraasi, 5 x *EvaGreen*® qPCR puhvrit, 12.5 mM MgCl₂'d, dNTP-sid, *EvaGreen*® värvainet, ROX värvainet, 0.2 µl praimerit F (10 pmol/µl), 0.2 µl praimerit R (10 pmol/µl), 1 µl genoomset DNA-d (10 ng/µl) ning ddH₂O-d lõppmahuni. Kõik reaktsioonid viidi läbi kolmes korduses kahe katsena. Kõikidele plaatidele lisati kvaliteedi kontrolliks üks referentsindiviid, kellel oli teadaolevalt 2 koopiat ning plaadisiseseks normaliseerimiseks kasutatav viljakate kontrollide DNA-de segu ja lahjendusrida. Kontrollide DNA segu valmistati 50-st juhuslikult valitud kontrollindiviidi DNA-st.

Tabel 3. *EvaGreen* qPCR-iga validatsiooniks valitud piirkonna asukoht, geeni nimi, F ja R praimerite järjestused ja kasutatud kogused.

CNV asukoht	Geen	Praimer F järjestus	Praimer R järjestus	Primeri kogus (µl)	
				R	F
11p13	<i>CAT</i>	AGGGTCTAAGTATTCCGTCTGC	CGGTTATCAGCATCCTTCAG	0.2	0.2
4q13.3	<i>ALB</i>	TGTTGCATGAGAAAACGCCA	GTCGCCTGTTACCAAGGATT	0.2	0.4

Amplifikatsioonireaktsioonid viidi läbi ABI Prism 7900HT Sequence Detection System (Applied Biosystems, USA) masinaga järgmistel tingimustel:

Algne denaturatsioon	95°C – 15 min	
Denaturatsioon	95°C – 15 s	} 40 tsüklit
Praimerite seondumine	60°C – 20 s	
Ekstensioon	72°C – 20s	

4.6 Andmete analüüs

TaqMan qPCR viidi läbi suhtelise kvantifitseerimise meetodil, *EvaGreen* qPCR absoluutse kvantifitseerimise meetodil ning andmete analüüsimiseks kasutati programme SDS 2.4, RQ Manager 1.2.1 (Applied Biosystems, USA) ja Microsoft Excel 2010. Koopiaarvu määramiseks normaliseeriti *TaqMan* qPCR-i puhul proovid referentsgeeni ehk *RNaseP* suhtes, *EvaGreen* qPCR-i puhul *ALB* suhtes.

5. TULEMUSED

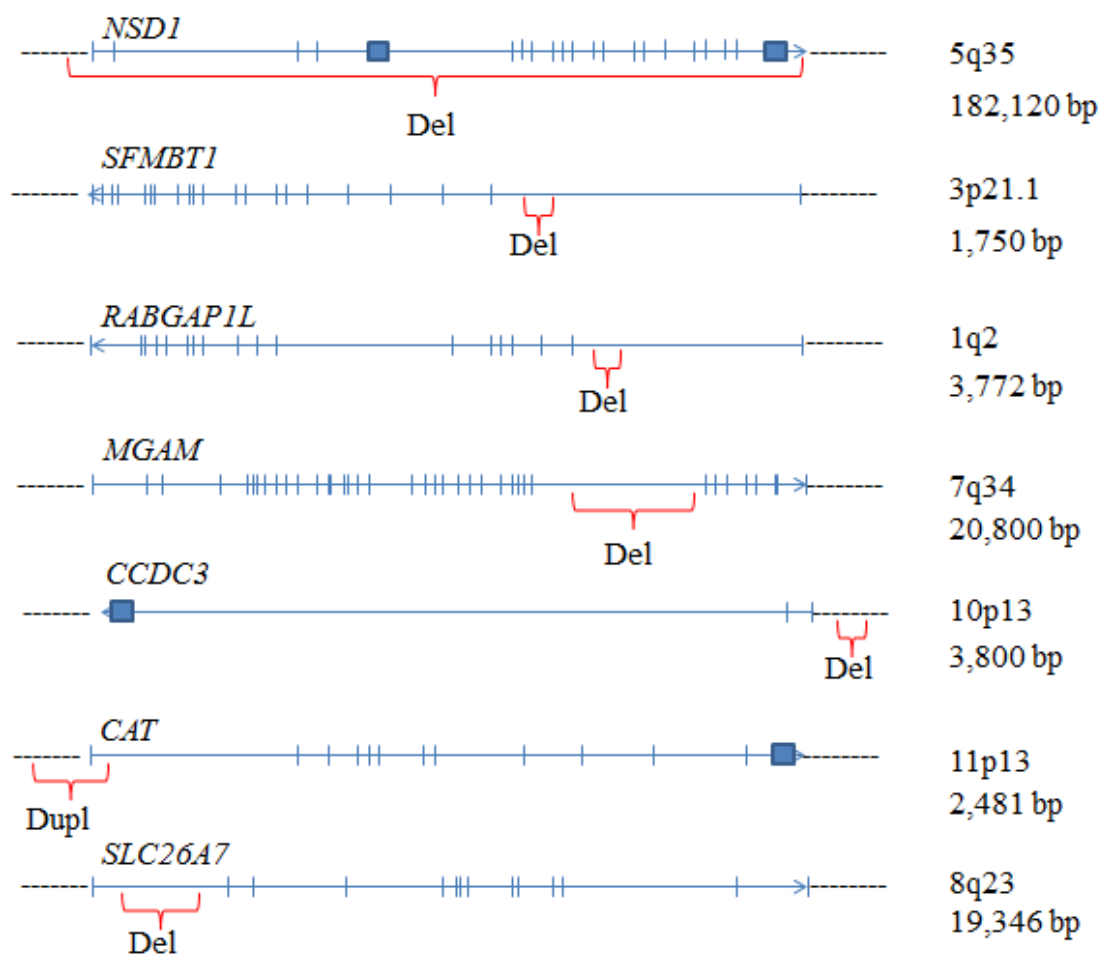
5.1. Tuvastatud CNV-de valik eksperimentaalseks validatsiooniks ja piirkondade iseloomustus

Varasemalt olid genotüpiseeritud Illumina HumanOmniExpress kiipidel erinevate raseduskomplikatsioonide ja normaalselt kulgenud raseduste platsenta DNA-d (joonis 2) ning vanemate verest eraldatud DNA-d. Valituks said: a) platsentade kiibipõhise CNV-de detektsiooni põhjal 4 piirkonda ning b) vanemate kiibipõhise CNV-de tuvastamise põhjal 3 piirkonda. Piirkondade eksperimentaalseks validatsiooniks valimisel oli kriteeriumiks, et CNV peab olema ennustatud vähemalt kahel indiviidil ühest patsiendigrupist. Lisaks arvestati CNV piirkondade poolt hõlmatud geenide kohta varasemalt avaldatud kirjandust ja nende ekspressiooni reproduktiivorganites (tabel 5). Valiti välja piirkonnad, mis võiksid olla bioloogiliselt olulised ning kõrgema esinemissagedusega erinevate rasedustüsistuste gruppides. Valitud piirkondasid iseloomustab tabel 4 ning CNV asukohti joonis 5.

Tabel 4. Validatsiooniks valitud piirkondade iseloomustus.

	Geeni nimi	Kromo- soom	CNV (Hg18)	algus CNV (Hg18)	lõpp CNV pikkus (bp)	CNV tüüp	Kandjate grupp
a)	<i>NSD1</i>	5q35	176,474,542	176,656,662	182,120	Deletsioon	2 LGA
							1 LGA,
	<i>MGAM</i>	7q34	141,419,097	141,439,888	20,800	Deletsioon	2 GD
	<i>RABGAP1L</i>	1q2	173,064,490	173,068,262	3,772	Deletsioon	2 PE
	<i>SFMBT1</i>	3p21.1	53,008,335	53,010,084	1,750	Deletsioon	2 LGA
b)	<i>SLC26A7</i>	8q23	92,344,944	92,364,290	19,346	Deletsioon	2 SGA
	<i>CAT</i>	11p13	34,414,806	34,417,287	2,481	Duplikatsioon	2 GD
	<i>CCDC3</i>	10p13	13,096,593	13,101,451	3,800	Deletsioon	2 SGA

a) platsentade ja b) vanemate kiibipõhise CNV detektsiooni põhjal valitud piirkonnad. *NSD1*: nuclear receptor binding SET domain protein 1; *MGAM*: maltase-glycoamylase; *RABGAP1L*: RAB GTPase activating protein 1-like isoform A; *SFMBT1*: Scm-like with four mbt domains 1; *SLC26A7*: solute carrier family 26, member 7 isoform b; *CAT*: catalase; *CCDC3*: coiled-coil domain containing 3. LGA: gestatsioonija kohta suur sünnikaal; SGA: gestatsioonija kohta väike sünnikaal; GD: gestatsioonidiabeet; PE: preeklampsia. Hg18: inimese genoomi järjestuse variant UCSC brauseris.



Joonis 5. CNV asukoht validatsiooniks valitud piirkondades. Horisontaalne katkendlik must joon kujutab geeni ümbritsevat ala; horisontaalne sinine noolega joon geeni, millel vertikaalsed sinised jooned ja kastid tähistavad eksoneid; punane sulg ennustatud CNV piirkonda. Paremal on toodud geeni asukoht ja pikkus. Del: deletsioon; dupl: duplikatsioon.

Tabel 5. Geenide avaldumine RNA ja valgu tasemel kudedes.

Geen	Ekspressioon kudedes RNA tasemel	Ekspressioon kudedes valgu tasemel
<i>SFMBT1</i>	Kõik koed, enim testistes	Kõik koed
<i>RABGAP1L</i>	Enim südamelihastes, ekspresseerub ka munasarjas, emakas, platsentas	Enim seedesüsteemis, uriinipõies, emakas
<i>MGAM</i>	Enim kaksteistsõrmiksooles, peensooles, neerudes	Enim seedesüsteemis, neerudes, luuüdis, kõrvalkilpnäärmes, ka munandimanuses, testises, platsentas
<i>CCDC3</i>	Enim sapipõies, söögitorus, nahas, rasvkoes, aga ka emakas, munasarjas	Pole infot
<i>NSD1</i>	Vähesel määral enamikes kudedes	Pole infot
<i>CAT</i>	Enim maksas, neerudes, rasvkoes, luuüdis, ka reproduktiivsüsteemis	Enim maksas, kaksteistsõrmiksooles, neerudes, ka eesnäärmes munasarjas, seemnepõiekestes
<i>SLC26A7</i>	Enim kilpnäärmes, reproduktiivsüsteemis vähe	Pole infot

5.2 Patsientide detailsem CNV-de profiil

NSD1 (*nuclear receptor binding SET domain protein 1*) geenis tuvastati deletsioon kahe LGA platsenta DNA-s. Kummalgi lapsel oli üks vanematest suure sünnikaaluga ning ülekaaluline. Varasemalt on antud geenis esinevaid mikroduplikatsioone seostatud Sotos sündroomiga, mille tunnusteks on mikrotsefaalia, väike kasv ja arengupeatetus. Antud geeni hõlmavad deletsioonid ja duplikatsioonid mõjutavad tugevalt keha arengut (Rosenfeld *et al.*, 2013).

SFMBT1 (*Scm-like with four mbt domains 1*) geeni haarav deletsioon oli ennustatud kahe LGA platsenta DNA-s. Üks isadest oli ülekaaluline, ning ema ja isa olid suure sünnikaaluga. *SFMBT1*

eksprimeerub kõrgelt trofoplasti rakkudes ja keskmiselt detsiidua rakkudes (Lin *et al.*, 2013). *Genetic Association Database* andmetel on seda geeni varasemalt seostatud kehamassiindeksi, kaalu ning pikkusega. *SFMBT1* represseerib embrüo arengus vajaminevaid geene ja arvatakse, et geen põhjustab hüpertensiooni teket, kuigi täpne mehhanism pole teada (Yang *et al.*, 2012).

RABGAP1L (*RAB GTPase activating protein 1-like isoform A*) geeni hõlmav homosügootne deletsioon oli ennustatud kahe preeklampsia patsiendi platsenta DNA-s. Gestatsiooniaeg oli mõlema raseduse korral normaalne. Mõlemal emal esines rasedusaegne hüpertensioon ning kõik vanemad olid ülekaalulised. Varasemalt on geeni seostatud hüpertensiooni (Oguri *et al.*, 2010) ja süsteemse luupus erütematoosusega. Luupusega naistel oli gestatsiooniaeg lühem, vastsündinu kaal väiksem ning sagedasemad preeklampsia juhud (Kim *et al.*, 2013). Luupusega seotud CNV oli sarnaste piiridega selles bakalaureusetöös ennustatud CNV-ga (vastavalt Chr1:173,064.5-173,068.3 ja Chr1:173,064.490-173,068.262, Hg18).

MGAM (*maltase-glycoamylase*) geeni haarav homosügootne deletsioon tuvastati ühel LGA ja kahel GD platsental. Üks ema ja kaks isa olid ülekaalulised, enamike vanemate sünnikaalud olid keskmisest suuremad. On näidatud, et *MGAM*-ilt transleeritakse maltaas-glükoamülaas, mis on vajalik tärglase lagundamiseks glükoosiks (Diaz-Sotomayor *et al.*, 2013). Geeni saaks kasutada ravimina tüüp II diabeedi ja kõrgekaalulisuse vastu (Ren *et al.*, 2011).

CCDC3 (*coiled-coil domain containing 3*) geenis ennustati deletsioon kahel SGA lapse vanemal. Üks ema oli alakaalus ning ülejäänud vanematel oli sünnikaal madal. Varem on näidatud, et *CCDC3* eksprimeerub vaskulaarsetes endoteelirakkudes ja rasvkoes (Azad *et al.*, 2014). *CCDC3* on konserveerunud funktsiooniga, hiirtel kõrgelt eksprimeerunud rasvkoes ning kuna skeletilihastes eksprimeerub vähe, siis on ilmselt seotud rasva või energia metabolismiga (Eberlein *et al.*, 2010).

CAT (*catalase*) geeni deletsioon ennustati kahel gestatsioonidiabeedi gruppi kuuluva lapse isal. Mõlemad isad olid ülekaalulised. *CAT* (katalaas) kodeerib ensüümi katalaasi, mis kaitseb keha oksüdatiivse stressi eest (Poggi *et al.*, 2012). Madal tase hõlbustab lihasrakkudes insuliini signalisatsiooni. Diabeedikute hulgas on geenis rohkem mutatsioone – vesinikperoksiid on üks diabeedi tekitajatest (Góth ja Nagy, 2012). Võrreldes normaalse platsentaga on GD platsentas kõrgem katalaasi ja teiste antioksidantide ekspressioon, nii et platsenta reageerib oksüdatiivsele

stressile vähem (Lappas et al., 2010). Samas on näidatud, et katalaasil on kaitsev efekt GD tekke eest (López-Tinoco *et al.*, 2013).

SLC26A7 (*solute carrier family 26, member 7 isoform b*) geeni deletsioon ennustati kahel SGA lapse emal. Mõlemad emad olid alakaalulised. Geen kuulub aniooni transportervalkude perekonda. Seda ekspresseeritakse hapet tootvas neeru tootmiskanalis, kus vahendab Cl⁻/HCO₃ vahetust (Sun ja Petrovic, 2008). *Slc26a7* kolokaliseerub *Slc26a2*-ga mao parietaalrakkudesse. Deletsiooni viimases on seostatud kasvuhäiretega (Xu *et al.*, 2009).

5.3 Kiibitulemuste eksperimentaalne validatsioon *TaqMan qPCR* meetodiga

Illumina HumanOmniExpress kiibipõhise CNV-de detektsiooni etapist valiti välja 7 CNV piirkonda, millega viidi läbi validatsioon, et kinnitada kiibikatses ennustatud ümberkorraldusi eksperimentaalselt. Validatsiooni etapp viidi läbi kiibikatses kasutatud valimiga (tabel 1).

Kiibikatsega ennustatud ning validatsiooni tulemuste võrdlus näitab, et validatsioonis kinnitati 7-st ennustatud piirkonnast 4 puhul vastavad CNV kandjad tõeseks (tabel 6). Lisaks tuvastati validatsioonikatsega erinevates piirkondades 1-7 uut CNV kandjat ehk valenegatiivset.

Tabel 6. Kiibipõhiselt ennustatud ja validatsiooniga tuvastatud CNV kandjate arv eksperimentaalselt valideerunud ja mittevalideerunud piirkondades.

	Piirkond	Kiibil ennustatud	Validatsioonis kinnitunud	Validatsioonis lisaks tuvastatud	Kokku
a)	<i>NSD1</i>	2	0	0	0
	<i>MGAM</i>	3	3	7	10
	<i>RABGAP1L</i>	2	2	7	9
	<i>SFMBT1</i>	2	2	4	6
b)	<i>SLC26A7</i>	2	0	0	0
	<i>CAT</i>	2	0	0	0
	<i>CCDC3</i>	2	2	1	3

a) platsentade ja b) vanemate kiibipõhise CNV detektsiooni põhjal valitud piirkonnad.

SLC26A7 geenis oli ennustatud deletsioon (CN=1) kahel hüpotroofia grupi naisel. *NSDI* geenis oli ennustatud deletsioon (CN=1) kahel makrosoomia grupi platsental. *CAT*-is oli ennustatud duplikatsioon (CN=4) kahel gestatsioonidiabeedi mehel. Nende piirkondade puhul validatsioon kiibi ennustust ei kinnitanud ning piirkondi edasi ei uuritud. *SLC26A7* ennustatud logBF väärtus oli piiripealne, ilmselt seetõttu ennustus ei kinnitunud. LogBF väärtus kinnitab, et CNV asub vastavas asukohas SNP andmete põhjal. Väärtustel 10-30 määratakse lühemaid CNV-sid ning võib esineda kuni 10% valepositiivseid tulemusi. LogBF alla 10 pole enamasti usaldusväärne (Colella *et al.*, 2007). *NSDI* puhul ei pruukinud ennustus tõene olla, kuna piirkonnas on palju inversioone. *CAT* piirkonna *TaqMan* proov oli valitud CNV ala äärde, mistõttu on võimalik, et CNV ei ulatu sinna.

CAT piirkonna puhul teostati lisaks validatsioon *EvaGreen qPCR*-iga, kinnitamaks, et *TaqMan qPCR* validatsiooni tulemused on tõesed ning kiibikatse tulemus valepositiivne. Katse kinnitas duplikatsiooni antud piirkonnas vaid ühes kahest ennustatud GD lapse isa DNA-s. Lisaks leiti hulgaliselt valenegatiivseid eri gruppidest. See tõestab, et kiibikatse tulemusi tuleb alati teise meetodiga kinnitada.

SFMBT1 piirkonnas leiti lisaks kahele makrosoomia platsentale deletsioonid kahel kontrollgrupi, ühel preeklampsia ja ühel gestatsioonidiabeedi grupi platsental (tabel 7).

Tabel 7. *SFMBT1* piirkonna validatsioonikatsete tulemused.

Platsentad	Arv	Kandjate arv CN=1	Kandjate sagedus %
Kontroll	8	2	25.0
Hüpotroofia	8	0	0.0
Preeklampsia	8	1	12.5
Makrosoomia	8	2	25.0
Gestatsioonidiabeet	8	1	12.5

CN: koopiaarv.

Validatsiooniga leiti *RABGAPIL* piirkonnas lisaks hulgaliselt hemisügootseid platsentasid – 2 kontrollgrupist, 1 hüpootroofia grupist, 3 makrosoomia ja 1 gestatsioonidiabeedi grupist (tabel 8).

Tabel 8. *RABGAPIL* piirkonna validatsioonikatsete tulemused.

Platsentad	Arv	Kandjate arv		Kandjate sagedus (%)	
		CN=1	CN=0	CN=1	CN=0
Kontroll	8	2	0	25.0	0.0
Hüpootroofia	8	1	0	12.5	0.0
Preeklampsia	8	0	2	0.0	25.0
Makrosoomia	8	3	0	37.5	0.0
Gestatsioonidiabeet	8	1	0	12.5	0.0

CN: koopiaarv.

MGAM piirkonna validatsiooniga selgus, et lisaks kiibikatsega ennustatule olid hemisügootsed 1 kontrollgrupi, 1 hüpootroofia ja 4 gestatsioonidiabeedi grupi platsentat (tabel 9).

Tabel 9. *MGAM* piirkonna validatsioonikatsete tulemused.

Platsentad	Arv	Kandjate arv		Kandjate sagedus (%)	
		CN=1	CN=0	CN=1	CN=0
Kontroll	8	1	0	12.5	0.0
Hüpootroofia	8	1	0	12.5	0.0
Preeklampsia	8	0	0	0.0	0.0
Makrosoomia	8	0	1	0.0	12.5
Gestatsioonidiabeet	8	4	2	50.0	25.0

CN: koopiaarv.

CCDC3 piirkonna validatsioonis kinnitus ennustatud deletsioon (tabel 10).

Tabel 10. CCDC3 piirkonna validatsioonikatsete tulemused.

	Arv	Kandjate arv CN=1	Kandjate sagedus (%) CN=1
Emad			
Kontroll	8	0	0
Hüpotroofia	8	1	12.5
Preeklampsia	8	0	0
Makrosoomia	8	0	0
Gestatsioonidiabeet	8	0	0
Isad			
Kontroll	8	0	0
Hüpotroofia	6	1	16.7
Preeklampsia	8	0	0
Makrosoomia	8	0	0
Gestatsioonidiabeet	8	0	0

5.4 Replikatsiooniuring laiendatud valimis

Valideeritud CNV piirkondadest valiti välja 4, millega viidi läbi replikatsiooniuring laiendatud valimis. Valimit laiendati huvipakkuvas grupis ja kontrollgrupis.

Validatsioonis oli *SFMBT1* geeni deletsiooniga kinnitatud 2 indiviidi nii makrosoomia kui ka kontrollgrupist, kuna valim oli väike, siis laiendati seda nendes gruppides. Lõppvalimis oli CNV esinemissagedus samuti kontrollgrupi ja makrosoomia gruppides võrdne (tabel 11). Hüpotees, et deletsioon antud piirkonnas makrosoomiat põhjustab, ei kinnitunud.

Tabel 11. *SFMBT1* replikatsiooniuuringu tulemused liidetuna validatsioonikatse tulemustega.

	Platsentad	Arv	Kandjate arv CN=1	Kandjate sagedus (%) CN=1	Alleelisagedus q (%)
a)	Kontroll	16	2	12.5	6.3
	LGA	16	2	12.5	6.3
b)	Kontroll	24	4	16.7	8.3
	LGA	24	4	16.7	8.3

a) replikatsiooniuuring; b) validatsiooni- ja replikatsiooniuuring. CN: koopiaarv; LGA: gestatsiooniaja kohta suure sünnikaaluga laps.

Validatsioonis kinnitus 0 koopiat *RABGAP1L* piirkonnas kahel preeklampsia platsental, seetõttu laiendati valimit kontrollgrupis ja preeklampsia grupis. Deletsiooni suhtes homosügootseid platsentasid rohkem juurde ei leitud, kuid hemisügootseid lisandus 10 kontroll- ja 5 preeklampsia grupist (tabel 12).

Tabel 12. *RABGAP1L* replikatsioonikatse tulemused liidetuna validatsioonikatse tulemustega.

	Platsentad	Arv	Kandjate arv		Kandjate sagedus (%)		Alleeli-sagedus q (%)	Statistiline olulisus	
			CN=1	CN=0	CN=1	CN=0		CN=1	CN=0
a)	Kontroll	48	10	0	21.2	0	10.4	0.409	1
	PE	37	5	0	13.5	0	6.8		
b)	Kontroll	57	12	0	21.7	0	10.5	0.284	0.187
	PE	44	5	2	11.4	4.5	10.2		

a) replikatsiooniuuring; b) validatsiooni- ja replikatsiooniuuring. Statistilise olulisuse määramiseks kasutati kahepoolset Fisheri täpset testi. CN: koopiaarv; PE: preeklampsia.

Hemisügootseid platsentasid esines igas grupis, mistõttu see suure tõenäosusega kindlat fenotüüpi ei määra. Käesoleva katse tulemuste põhjal pole PE platsentade DNA-s esinev homosügootne deletsioon *RABGAP1L* geenis usaldusväärsuse nivool $\alpha=0.05$ statistiliselt oluline

($P=0.187$). Kahjuks andmed ei tõesta ega lükka ümber, et homosügootne alleel tõstaks preeklampsia riski, seetõttu tuleks läbi viia rohkem katseid suuremas valimis.

RABGAP1L piirkonna deletsiooni kontrolliti kõigil emadel-isadel, kelle platsentad olid CNV kandjad. Kõikide puhul oli deletsioon vanematelt pärandunud. Mõlema homosügootse deletsiooni puhul oli tegemist hilise PE-ga.

Kuna *MGAM* geenis oli CNV sagedus rasedustüsistustest kõrgeim GD grupis, otsustati valimit laiendada selles ja kontrollgrupis. Lõppvalimis oli 0 koopia kandjaid kahes grupis võrdselt, 1 koopiat enam kontrollgrupi platsentadel (tabel 13).

Tabel 13. *MGAM* replikatsioonikatse tulemused liidetuna validatsioonikatse tulemustega.

	Platsentad	Arv	Kandjate arv		Kandjate sagedus (%)		Alleeli-sagedus q (%)	Statistiline olulisus	
			CN=1	CN=0	CN=1	CN=0		CN=1	CN=0
a)	Kontroll	43	1	2	2.3	4.7	5.8	0.202	1
	GD	42	4	2	9.5	4.8	9.5		
b)	Kontroll	51	2	2	3.9	3.9	5.9	0.092	0.436
	GD	50	7	4	16	8	15		

a) replikatsiooniuuring; b) validatsiooni- ja replikatsiooniuuring. Statistilise olulisuse määramiseks kasutati kahepoolset Fisheri täpset testi. CN: koopiaarv; GD: gestatsioonidiabeet.

Rikastatus gestatsioonidiabeedis küll kinnitus, kuid pole absoluutne. On võimalik, et deletsioon *MGAM* geenis on gestatsioonidiabeedi kõrvalpõhjuseks. Kontrollide hulgas on nii 0 kui 1 koopiaga indiviide. Tulemuste põhjal ei ole GD platsentade DNA-s esinev hemisügootne deletsioon *MGAM* geenis usaldusväarsuse nivool $\alpha=0.05$ statistiliselt oluline ($P=0.092$). Sama kehtib homosügootse deletsiooni puhul ($P=0.436$). Piirkonda tuleks uurida suuremas valimis.

Kuna kiibianalüüsil ennustati *CCDC3*-s deletsioon kahel hüpotroofia vanemal, laiendati valimit kontrollide ja hüpotroofia platsentade, naiste ja meeste puhul (tabel 14).

Tabel 13. *CCDC3* replikatsioonikatse tulemused liidetuna validatsiooni tulemustega.

		Grupp	Arv	Kandjate arv CN=1	Kandjate sagedus (%) CN=1	Alleeli- sagedus q (%)
Replikatsiooniuuring	Platsentad	Kontroll	64	2	3.1	1.6
		SGA	59	3	5.1	2.5
	Naised	Kontroll	55	0	0.0	0.0
		SGA	51	1	2.0	1.0
	Mehed	Kontroll	48	0	0.0	0.0
		SGA	41	3	7.3	3.7
Validatsiooni- ja replikatsiooniuuring	Platsentad	Kontroll	64	2	3.1	1.6
		SGA	59	3	5.1	2.5
	Naised	Kontroll	63	1	1.6	0.8
		SGA	59	1	1.7	0.8
	Mehed	Kontroll	56	0	0.0	0.0
		SGA	49	3	6.1	3.1

CN: koopiaarv; SGA: gestatsiooniaja kohta madal sünnikaaluga laps.

Deletsiooniga alleel pärandus platsentale ühelt SGA ja kontrollgrupi naiselt ning kahelt SGA grupi isalt. Tulemuste põhjal võib öelda, et deletsioon platsentas hüpotroofiaga seotud pole, kuna esinemissagedus on sarnane nii kontrollide kui ka hüpotroofia patsientide puhul.

Replikatsiooniuuringu tulemused näitavad, et kahes valitud piirkonnas (*RABGAP1L*, *MGAM*) esines huvipakkuvas grupis rikastatus võrreldes kontrollgrupiga.

6. ARUTELU

6.1. Metoodika usaldusväärsus

SNP mikrokiibi kasutamise eelisteks on analüüsi suur tihedus ja läbilaskevõime üle genoomi. Võrreldes aCGH-ga on meetod võimsam ja uurib genoomi laiemalt. Samas ei kata oligonukleotiidid igat genoomi osa võrdselt. SNP mikrokiibi analüüsi kasutatakse tavaliselt esmases sõeluuringus, et leida potentsiaalseid geneetilisi variatsioone, mida peab edaspidi eksperimentaalselt valideerima (Ho *et al.*, 2013).

Selles töös oli mikrokiibi tulemustest validatsiooniks valitud CNV piirkondade valikukriteeriumid karmid: CNV pidi olema ennustatud vähemalt kahe algoritmi poolt kolmest, vähemalt kahel indiviidil uuritavate grupist ning arvestati ka varasemat kirjandust geenide funktsioonide kohta ja ekspressiooni reproduktiivorganites. Neljas valideerunud piirkonnas valepositiivseid proove ei leitud, küll aga tuvastati piirkonna kohta 1-7 valenegatiivset. Selle tagasid CNV-de tuvastamise ranged kriteeriumid. Siiski leidsid mõned CNV-d, mis ei kinnitunud. See võib olla tingitud keerukast genoomsest struktuurist. Samuti tähendab see seda, et mikrokiibi põhised tulemused tuleks alati alternatiivse meetodiga üle valideerida.

Piirkonnad, mis ei valideerunud, hõlmasid *NDS1*, *CAT* ja *SLC26A7* geene. Esimese piirkonna puhul võis ennustus ebatõene olla, kuna seal asub hulgaliselt inversioone. *CAT* piirkonna puhul olid praimerid CNV ala äärde valitud ning on võimalik, et mikrokiibi ennustatud duplikatsiooni pikkus oli tegelikkuses lühem. *SLC26A7* piirkonna deletsioonil oli logBF piiripealne, mistõttu oli ennustus ilmselt vale.

qPCR-i eeliseks on madal hind, vajamineva varustuse lai kättesaadavus ja kvantifitseerimise ja analüüsi ühised meetodid. Samuti pole vaja teada ka väga palju genoomset infot. Puudusteks on see, et on võimalik analüüsida vähe geene korraga ning endogeenseid kontrolle peab hoolikalt valima (Fassbinder-Orth, 2014).

Antud töös tõestati, et *TaqMan* qPCR oli usaldusväärne vahend SNP mikrokiibi tulemuste kontrollimiseks. Esmalt teostati *CAT* piirkonna validatsioon *TaqMan* qPCR meetodiga, mis mikrokiibi tulemusi ei kinnitanud. Seejärel teostati validatsioon ka *EvaGreen* qPCR-iga. Erinevalt esimesest meetodist, kus detekteeritakse fluorestsentssignaali tugevnemist *TaqMan* proovi lagunedes, seondub *EvaGreen* värv mittespetsiifiliselt kaheaheelalisele DNA-le (Arya *et*

al., 2005; Mao *et al.*, 2007). *EvaGreen* qPCR-iga kinnitus deletsioon vaid ühes gestatsioonidiabeedi mehe DNA-s kahest. See tõestab, et *TaqMan* qPCR-i tulemused on usaldusväärsed.

6.2 Uuritud CNV-d ja nende seos haigustega

Käesoleva töö raames viidi läbi katsed, et kinnitada varasemalt teostatud SNP kiibikatse tulemusi ning teostada vajadusel replikatsioonikatseid, et uurida valitud piirkondades esinevate koopiaarvu varieeruvuste rolli seoses erinevate rasedusaegsete komplikatsioonidega.

CNV-del on erinevaid fenotüüpilisi efekte, seoses nii kohastumise kui letaalsusega. Patogeensed CNV-d on seotud erinevate haigustega ning esineb ka selliseid, mis haigestumist soodustavad (Zarrei *et al.*, 2015). Haiguseid tekitavad CNV-d on enamasti üle miljoni aluspaari pikad ning hõlmavad geenirikkaid alasid (Savage *et al.*, 2011). Varasemalt on teostatud väga vähe uuringuid seoses CNV-dega, mis võiksid olla seotud rasedusaegsete komplikatsioonidega. Uuritud on vaid SGA-d ja PE-d mõjutavaid koopiaarvu variatsioone. Selles töös seostati preeklampsia deletsiooni *RABGAP1L* geenis ja hüpotroofiaga deletsioone *CCDC3* ja *SLC26A7* geenides. Lisaks uuriti potentsiaalselt LGA ja GD-ga seotud CNV-sid - vastavalt deletsioonid *NSD1*, *SFMBT1* geenides ja deletsioon *MGAM* ning duplikatsioon *CAT* geenides.

Kahes uuritud piirkonnas, mis valideerusid, oli deletsioonikandjaid uuritavas grupis võrdselt. Need piirkonnad olid *SFMBT1* ja *CCDC3*. Tulemuste põhjal võib väita, et deletsioonid nendes piirkondades ei ole seotud rasedusaegsete komplikatsioonidega.

RABGAP1L geeni on varasemalt kirjanduses seostatud preeklampsia ja väikese sünnikaaluga (Kim *et al.*, 2013) ning on teada, et see ekspresseerub ka reproduktiivorganites. Validatsiooni- ja replikatsioonikatsetega leiti 17 hemisügootset platsenta DNA-d. On alust arvata, et heterosügootne deletsioon preeklampsia seotud pole, kuna seda esines kontrollgrupi platsentadel kaks korda rohkem. Peale kiibikatses ennustatud kahe platsenta DNA, *RABGAP1L* geenis homosügootse deletsiooni kandjaid validatsiooni- ega replikatsioonikatses ei lisandunud. Oluline oleks teostada replikatsioonikatse suurema valimiga, et kindlalt väita, kas deletsioon antud piirkonnas mängib rolli preeklampsia tekkes. Samas on keeruline leida ekstreemseid

preeklampsia juhtusid ja platsentamaterjali. Huvitavaks asjaoluks on see, et samasuguste ennustatud piiridega deletsiooni *RABGAPIL* geenis on varem seostatud luupusega. Luupusega rasedatel naistel on suurem tõenäosus preeklampsia. Lisaks esines selles uuringus homosügootset deletsiooni luupusega naistel kaks korda enam (vastavalt 9.2% ja 5.5%) (Kim *et al.*, 2013).

MGAM geeni produkt töötab tähtsuse lagundamises glükoosiks (Diaz-Sotomayor *et al.*, 2013) ning ekspresseerub samuti reproduktiivorganites. *MGAM* piirkonna homosügootset deletsiooni ennustas SNP mikrokiibi katse kahel GD ja ühel LGA platsenta DNA-l. Validatsiooniga tuvastati lisaks neile 6 hemisügooti ning replikatsioonikatses 4 hemisügootset ja 2 homosügootset deletsiooni. Heterosügootsete deletsioonikandjate hulgas võib näha rikastust gestatsioonidiabeedi grupis. Tulemus pole statistiliselt oluline, aga see võib olla tingitud väikesest valimist. Vaja oleks läbi viia katseid suurema valimiga, et kinnitada või ümber lükata *MGAM* geenis esineva deletsiooni mõju gestatsioonidiabeedi tekkes.

Kahe piirkonna, *RABGAPIL* ja *CCDC3* puhul uuriti *qPCR*-iga nii platsentade DNA-sid kui ka nende vanemate omasid. Triosid on raseduse puhul oluline uurida, kuna komplikatsiooni teket võivad soodustada kõigi DNA-s asuvad struktuursed ümberkorraldused ning ainult kogu pere geneetilise materjali uurimisel saab täielikku infot selle kohta, kas CNV on päritud või tekkinud *de novo*.

Käesoleva bakalaureusetöö eksperimentaalse osa põhjal ei ole võimalik väita, et mõni uuritud CNV-dest kindlalt rasedusaegseid komplikatsioone põhjustaks. Samas leidub *RABGAPIL* ja *MGAM* piirkondade puhul CNV-de rikastus patsientide grupis võrreldes kontrollidega, mida võiks edasi uurida suurema valimiga.

Käesoleva töö tähtsus paigutub selles, et uuriti esmakordselt trio (ema-isa-platsenta) DNA materjalis CNV piirkondi, mis võiksid olla seotud selliste rasedusaegsete komplikatsioonidega nagu gestatsioonidiabeet, preeklampsia, gestatsiooniaja kohta väikene ja suur sünnikaal. Varasemalt on CNV-dega seoses neist uuritud vaid preeklampsiat ja SGA-d. Rasedusaegsete komplikatsioonide põhjused on nii keskkonnast mõjutatud kui ka geneetilised. Geneetilist poolt ei tunta piisavalt hästi. Samas haigestub nendesse hulgaliselt Eesti rasedaid naisi. Seetõttu on oluline uurida raseduse teises pooles esinevate komplikatsioonidega seotud CNV-sid, mõistmaks

ja leidmaks nendele lahendusi. Lähiperspektiiviks oleks laiendada *MGAM* ja *RABGAP1L* piirkondade puhul valimit, kinnitamaks või ümber lükkamaks deletsioonide rikastatus nendes.

KOKKUVÕTE

Varasemalt on ilmunud väga vähe uurimusi, mis uuriks koopiaarvu varieeruvusi seoses rasedusaegsete komplikatsioonidega. Teada on küll keskkonnafaktorid, mis neid haiguseid põhjustavad, kuid samas on olulised ka geneetilised faktorid. Seetõttu on vaja rasedustüsistuste paremaks mõistmiseks teostada rohkem uurimusi, leidmaks nendega seotud CNV-sid.

Töö eesmärgiks oli varasemalt teostatud SNP mikrokiibi katsete tulemuste kinnitamine ja täiendavate replikatsioonikatsete läbiviimine, tõestamaks või ümber lükkamaks valitud piirkondades asuvate CNV-de olemasolu ja seost erinevate rasedusaegsete komplikatsioonidega. Katsed viidi läbi, kasutades *TaqMan* ja *EvaGreen qPCR*-i.

Validatsioonikatsetega kinnitus uuritud 7-st piirkonnast 4 ning nende uurimiseks teostati lisakatseid. *SFMBT1* ja *CCDC3* piirkondades esines vastav CNV võrdselt uuritava ja kontrollgrupi vahel. *RABGAP1L* ja *MGAM* piirkondades esines kerge rikastatus uuritavas grupis, mis pole aga statistiliselt oluline. Seega ei saa antud töö põhjal väita, et mõni CNV piirkond oleks kindlalt seotud mõne rasedusaegse komplikatsiooniga.

Käesolevas töös kasutatud valim oli väike, mis võib ka olla põhjuseks, miks ühtegi CNV-d kindlalt komplikatsioonidega seostada ei saa. Piirkondades, kus esines CNV-de rikastatus rasedustüsistuse grupis võrreldes normaalse rasedusega, annaks valimi suurendamine selgust, kas ümberkorraldatud geenid võiksid põhjustada eelsoodumust hilisteks rasedusaegseteks komplikatsioonideks.

Antud töö on esimene, mis uurib raseduse hiliskäigus esinevate rasedustüsistustega seotud koopiaarvu varieeruvusi. Lisaks on tegemist platsenta-ema-isa trio uuringuga, mis on oluline mõistmaks, kuidas CNV-d põlvkondade vahel päranduvad.

Copy number variations in complicated pregnancies

Anna-Maria Himma

SUMMARY

Several pregnancy complications can occur during late pregnancy. These can be: 1) gestational diabetes - pregnant women who are carbohydrate intolerant (Ferrara *et al.*, 2014); 2) preeclampsia, which symptoms are high blood pressure and high levels of protein in urine (Zhao *et al.*, 2012); 3) newborns who are born too large for gestational age and 4) too small for gestational age. These complications might have a greater effect on newborns, for example gestational diabetes may cause the baby to be born with adverse birth defects and to develop glucose intolerance and obesity later in life (Zhang *et al.*, 2014).

Copy number variations are at least 50 base pairs long insertions, deletions, duplications, inversions and translocations (Weischenfeldt *et al.*, 2013). While autosomes usually consist of two copies of each segment, the genesis of copy number variations alter that. Deletion means that we have zero or one copy of the original segment left, duplication that we have 3 or more copies (Savage *et al.*, 2011).

Whole genome SNP microarray has been used previously on 40 placentas and their parents DNAs to perform a whole-genome screening of CNVs in late pregnancy complications. Same DNA samples were used in validation experiments as well. REPROMETA cohort consists of DNA samples from Estonian pregnant women, their partners and placenta. The aim of this study was to confirm the results of the array experiment and to find out if copy number variations were the reason for any of the pregnancy complications mentioned above. In this study 4 out of 7 CNV areas chosen from SNP microarray results were validated, using *TaqMan qPCR* and REPROMETA cohort. This confirms that it is necessary to validate SNP microarray results, because 3 CNV areas were falsely predicted.

CNV-s that were confirmed by validation with *TaqMan qPCR* were studied further. Deletions encompassing *SFMBT1* ja *CCDC3* genes were not related with any of the complications as the CNV frequency was similar between both the study and control group. The other two regions, *MGAM* and *RABGAP1L*, had more deletion carriers in the study group. The results were not statistically relevant, therefore we cannot confirm if these CNVs are related to any of the pregnancy complications. The number of samples used in this experiment was relatively small. To confirm or disprove the connection between these CNVs and pregnancy complications a new experiment with a larger group of samples should be carried out.

In this study DNA-s from both mothers, fathers and placentas were used for the first time to find out connections between chosen CNVs and pregnancy complications such as gestational diabetes, preeclampsia, small and large for gestational age babies. Pregnancy complications are known to be caused by both environmental and genetical factors, but genetic ones have not been studied enough. As many Estonian women are affected by these diseases during pregnancy, it is important to understand and find solutions to those complications.

KIRJANDUSE LOETELU

Ahlsson, F., Åkerud, H., Schijven, D., Olivier, J., Sundström-Poromaa, I. (2015). Gene Expression in Placentas From Nondiabetic Women Giving Birth to Large for Gestational Age Infants. *Reprod Sci*.

Alberico, S., Montico, M., Barresi, V., Monasta, L., Businelli, C., Soini, V., Erenbourg, A., Ronfani, L., Maso, G. (2014). The role of gestational diabetes, pre-pregnancy body mass index and gestational weight gain on the risk of newborn macrosomia: results from a prospective multicentre study. *BMC Pregnancy and Childbirth*, 14: 23.

Allvee, K., Karro, H. (2014). Eesti meditsiiniline sünniregister 1992-2013. Eesti abordiregister 1996-2013. Tervise Arengu Instituut. Tallinn.

Armour, J. A., Sismani, C., Patsalis, P. C., Cross, G. (2000). Measurement of locus copy number by hybridisation with amplifiable probes. *Nucleic Acids Res*. 28(2):605-9.

Arya, M., Shergill, I. S., Williamson, M., Gommersall, L., Arya, N., Patel, H. R. H. (2005). Basic principles of real-time quantitative PCR. *Expert Rev. Mol. Diagn*. 5(2), 209-219.

Azad, A. K., Chakrabarti, S., Xu, Z., Davidge, S. T., Fu, Y. (2014). Coiled-coil domain containing 3 (*CCDC3*) represses tumor necrosis factor- α /nuclear factor κ B-induced endothelial inflammation. *Cell Signal*. 26(12):2793-800.

Beltcheva, O., Boyadzhieva, M., Angelova, O., Mitev, V., Kaneva, R., Atanasova, I. (2014). The rs266729 single-nucleotide polymorphism in the adiponectin gene shows association with gestational diabetes. *Arch Gynecol Obstet*. 289(4):743-8.

Buckley, B. S., Harreiter, J., Damm, P., Chico, A., Simmons, D., Vellinga, A., Dunne, F. (2012). Gestational diabetes mellitus in Europe: prevalence, current screening practice and barriers to screening. A review. *Diabet. Med*. 29, 844–854.

Bustamante, M., Danileviciute, A., Espinosa, A., Gonzalez, J. R., Subirana, I., Cordier, S., Chevrier, C., Chatzi, L., Grazuleviciene, R., Sunyer, J., Ibarluzea, J., Ballester, F., Villanueva, C. M., Nieuwenhuisen, M., Estivill, X., Kogevinas, M. (2012). Influence of fetal glutathione S-

transferase copy number variants on adverse reproductive outcomes. *BJOG*. Volume 119, Issue 9. Pages 1141–1146

Canton, A. P., Costa, S. S., Rodrigues T. C., Bertola, D. R., Malaquias, A. C., Correa, F. A., Arnhold, I. J., Rosenberg, C., Jorge A. A. (2014). Genome-wide screening of copy number variants in children born small for gestational age reveals several candidate genes involved in growth pathways. *Eur J Endocrinol*. 171(2): 253-262.

Carreira, P. E., Richardson, S. R., Faulkner, G. J. (2014). L1 retrotransposons, cancer stem cells and oncogenesis. *FEBS J*. 281(1):63-73.

Chawla, R., Badon, S. E., Rangarajan, J., Reisetter, A. C., Armstrong, L. L., Lowe, L. P., Urbanek, M., Metzger, B. E., Hayes, M. G., Scholtens, D. M., Lowe, W. L. Jr. (2014). Genetic risk score for prediction of newborn adiposity and large-for-gestational-age birth. *J Clin Endocrinol Metab*. 99(11):E2377-86.

Colella, S., Yau, C., Taylor, J. M., Mirza, G., Butler, H., Clouston, P., Bassett, A. S., Seller, A., Holmes, C. C., Ragoussis, J. (2007). QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res*. 35(6): 2013–2025.

Diaz-Sotomayor, M., Quezada-Calvillo, R., Avery, S. E., Chacko, S. K., Yan, L. K., Lin, A. H., Ao, Z. H., Hamaker, B. R., Nichols, B. L. (2013). Maltase-glucoamylase modulates gluconeogenesis and sucrase-isomaltase dominates starch digestion glucogenesis. *J Pediatr Gastroenterol Nutr*. 57(6):704-12.

Dittwald, P., Gambin, T., Szafranski, P., Li, J., Amato, S., Divon, M. Y., Rojas, L. X. R., Elton, L. E., Scott, D. A., Schaaf, C. P., Torres-Martinez, W., Stevens, A. K., Rosenfeld, J. A., Agadi, S., Francis, D., Kang, S-H. L., Breman, A., Lalani, S. R., Bacino, C. A., Bi, W., Milosavljevic, A., Beaudet, A. L., Patel, A., Shaw, C. A., Lupski, J. R., Gambin, A., Cheung, S. W., Stankiewicz, P. (2013). NAHR-mediated copy-number variants in a clinical population: Mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res*. 23(9): 1395–1409.

- Dumas, L., Kim, Y. H., Karimpour-Fard, A., Cox, M., Hopkins, J., Pollack, J. R., Sikela, J. M. (2007). Gene copy number variation spanning 60 million years of human and primate evolution. *Genome Res.* 17(9): 1266–1277.
- Eberlein, A., Kalbe, C., Goldammer, T., Brunner, R. M., Kuehn, C., Weikard, R. (2010). Analysis of structure and gene expression of bovine *CCDC3* gene indicates a function in fat metabolism. *Comp Biochem Physiol B Biochem Mol Biol.* 156(1):19-25.
- Fassbinder-Orth, C. A. (2014). Methods for quantifying gene expression in ecoimmunology: from qPCR to RNA-Seq. *Integr Comp Biol.* 54(3):396-406.
- Ferrara, A., Hedderson, M. M., Albright, C. L., Brown, S. D., Ehrlich, S. F., Caan, B. J., Sternfeld, B., Gordon, N. P., Schmittiel, J. A., Gunderson, E. P., Mevi, A. A., Tsai, A., Ching, J., Crites, Y., Quesenberry Jr., C. P. (2014). A pragmatic cluster randomized clinical trial of diabetes prevention strategies for women with gestational diabetes: design and rational of the Gestational Diabetes' Effects on Moms (GEM) study. *BMC Pregnancy and Childbirth.* 14: 21.
- Girirajan, S., Campbell, C. D., Eichler, E. E. (2011). Human Copy Number Variation and Complex Genetic Disease. *Annu. Rev. Genet.* 45: 203-226.
- Góth, L., Nagy, T. (2012). *Acatlasemia* and diabetes mellitus. *Arch Biochem Biophys.* 525(2):195-200.
- Harmon, Q. E., Engel, S. M., Wu, M. C., Moran, T. M., Luo, J., Stuebe, A. M., Avery, C. L., Olshan, A. F. (2014). Polymorphisms in inflammatory genes are associated with term small for gestational age and preeclampsia. *Am J Reprod Immunol.* 71(5):472-84.
- Hinkle, S. N., Albert, P.S., Mendola, P., Sjaarda, L. A., Boghossian, N. S., Yeung, E., Laughon, S. K. (2014). Differences in risk factors for incident and recurrent small-for-gestational-age birthweight: a hospital-based cohort study. *BJOG.* 121: 1080-1089.
- Ho, C. C., Mun, K. S., Naidu, R. (2013). SNP array technology: an array of hope in breast cancer research. *Malays J Pathol.* 35(1):33-43.

- Hodgetts, V. A., Morris, R. K., Francis, A., Gardosi, J., Ismail, K. M. (2015). Effectiveness of folic acid supplementation in pregnancy on reducing the risk of small-for-gestational age neonates: a population study, systematic review and meta-analysis. *BJOG*. 2015 Mar;122(4):478-90.
- Iacobucci, I., Lonetti, A., Papayannidis, C., Martinelli, G. Curr. (2013). Use of Single Nucleotide Polymorphism Array Technology to Improve the Identification of Chromosomal Lesions in Leukemia. *Cancer Drug Targets*. 13(7):791-810.
- Kallioniemi, A., Kallioniemi, O-P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., Pinkel, D. (1992). Comparative Genomic Hybridization for Molecular Cytogenetic Analysis of Solid Tumors. *Science*. 258(5083):818-21.
- Kellogg, D. E., Sninsky, J. J., Kwok, S. (1990). Quantitation of HIV-1 proviral DNA relative to cellular DNA by the polymerase chain reaction. *Anal Biochem*. 189(2):202-8.
- Khong, S. L., Kane, S. C., Brennecke, S. P., Costa, F. da S. (2015). First-Trimester Uterine Artery Doppler Analysis in the Prediction of Later Pregnancy Complications. *Dis Markers*. 2015: 679730.
- Kidd, J. M., Cooper G. M., Donahue, W. F., Hayden, H. S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., Haugen, E., Zerr, T., Yamada, N. A., Tsang, P., Newman, T. L., Tüzün, E., Cheng, Z., Ebling, H. M., Tusneem, N., David, R., Gillett, W., Phelps, K. A., Weaver, M., Saranga, D., Brand, A., Tao, W., Gustafson, E., McKernan, K., Chen, L., Malig, M., Smith, J. D., Korn, J. M., McCarroll, S. A., Altshuler, D. A., Peiffer, D. A., Dorschner, M., Stamatoyannopoulos, J., Schwartz, D., Nickerson, D. A., Mullikin, J. C., Wilson, R. K., Bruhn, L., Olson, M. V., Kaul, R., Smith, D. R., Eichler, E. E. (2008). Mapping and sequencing of structural variation from eight human genomes. *Nature*. 453(7191):56-64.
- Kim, J. H., Jung, S. H., Bae, J. S., Lee, H. S., Yim, S. H., Park, S. Y., Bang, S. Y., Hu, H. J., Shin, H. D., Bae, S. C., Chung, Y. J. (2013). Deletion variants of *RABGAP1L*, 10q21.3, and C4 are associated with the risk of systemic lupus erythematosus in Korean women. *Arthritis Rheum*. 65(4):1055-63.

Korn, J. M., Kuruvilla, F. G., McCarroll, S. A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P. J., Darvishi, K., Lee, C., Nizzari, M. M., Gabriel, S. B., Purcell, S., Daly, M. J., Altshuler, D. (2008). Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet.* 40(10): 1253–1260.

Kutsar, D. (2012). *Living Conditions in Estonia Five Years Later*. Tartu.

Lambert, G., Brichant, J. F., Hartstein, G., Bonhomme, V., Dewandre, P. Y. (2014). Preeclampsia: an update. *Acta Anaesthesiol Belg.* 65(4):137-49.

Lappas, M., Mitton, A., Permezel, M. (2010). In response to oxidative stress, the expression of inflammatory cytokines and antioxidant enzymes are impaired in placenta, but not adipose tissue, of women with gestational diabetes. *J Endocrinol.* 204(1):75-84.

Li, W., Olivier, M. (2013). Current analysis platforms and methods for detecting copy number variation. *Physiol Genomics.* 45(1): 1–16.

Lin, S., Shen, H., Li, J. L., Tang, S., Gu, Y., Chen, Z., Hu, C., Rice, J. C., Lu, J., Wu, L. (2013). Proteomic and functional analyses reveal the role of chromatin reader *SFMBT1* in regulating epigenetic silencing and the myogenic gene program. *J Biol Chem.* 288(9):6238-47.

Liu, P., Carvalho, C. M. B., Hastings, P. J., Lupski, J. (2012). Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev.* 22(3): 211–220.

Liu, G. E., Xu, L., Huang, K. S. (2014). Recent Advances in Studying of Copy Number Variation and Gene Expression. *Gene Expression to Genetical Genomics.* 7 1-5.

López-Tinoco, C., Roca, M., García-Valero, A., Murri, M., Tinahones, F. J., Segundo, C., Bartha, J. L., Aguilar-Diosdado, M. (2013). Oxidative stress and antioxidant status in patients with late-onset gestational diabetes mellitus. *Acta Diabetol.* 50(2):201-8.

Luke, B., Brown, M. B. (2007). Elevated risks of pregnancy complications and adverse outcomes with increasing maternal age. *Hum. Reprod.* 22 (5): 1264-1272.

Mao, F., Leung, W-Y., Xin, X. (2007). Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnology.* 7:76.

- Männik, K., Parkel, S., Palta, P., Zilina, O., Puusepp, H., Esko, T., Mägi, R., Nõukas, M., Veidenberg, A., Nelis, M., Metspalu, A., Remm, M., Ounap, K., Kurg, A. (2011). A parallel SNP array study of genomic aberrations associated with mental retardation in patients and general population in Estonia. *Eur J Med Genet.* 54(2):136-43.
- Oguri, M., Kato, K., Yokoi, K., Yoshida, T., Watanabe, S., Metoki, N., Yoshida, H., Satoh, K., Aoyagi, Y., Nozawa, Y., Yamada, Y. (2010). Assessment of a polymorphism of *SDK1* with hypertension in Japanese Individuals. *Am J Hypertens.* 23(1):70-7.
- Poggi, C., Giusti, B., Vestri, A., Pasquini, E., Abbate, R., Dani, C. (2012). Genetic polymorphisms of antioxidant enzymes in preterm infants. *J Matern Fetal Neonatal Med. Suppl* 4:131-4.
- Redman, C. W. G. (2011). Preeclampsia: A multi-stress disorder. *Rev Med Interne.* 32 Suppl 1:S41-4.
- Ren, L., Qin, X., Cao, X., Wang, L., Bai, F., Bai, G., Shen, Y. (2011). Structural insight into substrate specificity of human intestinal maltase-glucoamylase. *Protein Cell.* 2(10):827-36.
- Rosenfeld, J. A., Kim, K. H., Angle, B., Troxell, R., Gorski, J. L., Westemeyer, M., Frydman, M., Senturias, Y., Earl, D., Torchia, B., Schultz, R. A., Ellison, J. W., Tsuchiya, K., Zimmerman, S., Smolarek, T. A., Ballif, B. C., Shaffer, L. G. (2013). Further Evidence of Contrasting Phenotypes Caused by Reciprocal Deletions and Duplications: Duplication of *NSD1* Causes Growth Retardation and Microcephaly. *Mol Syndromol.* 3(6):247-54.
- Rossi, A. C., Mullin, P., Prefumo, F. (2013). Prevention, Management, and Outcomes of Macrosomia: A Systematic Review of Literature and Meta-analysis. *Obstet Gynecol Surv.* 68(10):702-9.
- Sahoo, T., Wang, J-C., Elnaggar, M. M., Sanchez-Lara, P., Ross, L. P., Mahon, L. W., Hafezi, K., Deming, A., Hinman, L., Bruno, Y., Bartley, J. A., Liehr, T., Anguiano, A., Jones, M. (2015). Concurrent triplication and uniparental isodisomy: evidence for microhomology-mediated break-induced replication model for genomic rearrangements. *Eur J Hum Genet.* 23(1):61-6.

- Saenger, P., Reiter, E. (2012). Genetic factors associated with small for gestational age birth and the use of human growth hormone in treating the disorder. *Int J Pediatr Endocrinol.* 2012(1):12.
- Savage, M. S., Mourad, M. J., Wapner, R. J. (2011). Evolving applications of microarray analysis in prenatal diagnosis. *Curr Opin Obstet Gynecol.* 23: 103-108.
- Schouten, J. P., McElgunn, C. J., Waaijer, R., Zwiijnenburg, D., Diepvens, F., Pals, G. (2002). Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* 30(12):e57.
- Strassberg, M., Fruhman, G., den Veyver, I. B. V. (2011). Copy-number changes in prenatal diagnosis. *Expert Rev. Mol. Diagn.* 11(6), 579-592.
- Sun, X., Petrovic, S. (2008). Increased acid load and deletion of AE1 increase *Slc26a7* expression. *Nephron Physiol.* 109(3):p29-35.
- Uusküla, L., Männik, J., Rull, K., Minajeva, A., Kõks, S., Vaas, P., Teesalu, P., Reimand, J., Laan, M. (2012). Mid-gestational gene expression profile in placenta and link to pregnancy complications. *PLoS One.*
- Vanderweyer, G., Kooy, R. F. (2013). Detection and interpretation of genomic structural variation in health and disease. *Expert Rev Mol Diagn.* 13(1):61-82.
- Vora, N., Bianchi, D. W. (2009). Genetic considerations in the prenatal diagnosis of overgrowth syndromes. *Prenat Diagn.* 29(10):923-9.
- Waters, C. A., Strande, N. T., Wyatt, D. W., Pryor, D. W., Ramsden, D. A. (2014). Nonhomologous end joining: a good solution for bad ends. *DNA Repair (Amst).* 17:39-51.
- Weischenfeldt, J., Symmons, O., Spitz, F., Korbel, J. O. (2013). Phenotypic impact of genomic structural variation: insights from and for human disease. *Nature Reviews Genetics.* 14, 125-138.
- Winchester, L., Yau, C., Ragoussis, J. (2009). Comparing CNV detection methods for SNP arrays. *Brief Funct Genomic Proteomic.* 8(5):353-66.

Woelkers, D., Barton, J., Dadelszen, P., Sibai, B. (2015). The revised 2013 ACOG definitions of hypertensive disorders of pregnancy significantly increase the diagnostic prevalence of preeclampsia. *Pregnancy Hypertens.* 2015 Jan;5(1):38.

Xu, J., Song, P., Nakamura, S., Miller, M., Barone, S., Alper, S. L., Riederer, B., Bonhagen, J., Arend, L. J., Amlal, H., Seidler, U., Soleimani, M. (2009). Deletion of the Chloride Transporter *Slc26a7* Causes Distal Renal Tubular Acidosis and Impairs Gastric Acid Secretion. *J Biol Chem.* 284(43): 29470–29479.

Yang, H-C., Liang, Y-J., Chen, J-W., Chiang, K-M., Chung, C-M., Ho, H-Y., Ting, C-T., Lin, T-H., Sheu, S-H., Tsai, W-C., Chen, J-H., Leu, H-B., Yin, W-H., Chiu, T-Y., Chern, C-I., Lin, S-J., Tomlinson, B., Guo, Y., Sham, P. C., Cherny, S. S., Lam, T. H., Thomas, G. H., Pan, W. H. (2012). Identification of *IGF1*, *SLC4A4*, *WWOX*, and *SFMBT1* as Hypertension Susceptibility Genes in Han Chinese with a Genome-Wide Gene-Based Association Study. *PLoS One.* 7(3): e32907.

Zarrei, M., MacDonald, J. R., Merico, D., Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat Rev Genet.* 2015 Mar;16(3):172-83.

Zhang, F., Gu, W., Hurles, M. E., Lupski, J. R. (2009). Copy Number Variation in Human Health, Disease, and Evolution. *Annu. Rev. Genomics Hum. Genet.* 10: 451-48.

Zhang, C., Tobias, D. K., Chavarro, J. E., Bao, W., Wang, D., Ley, S. H., Hu, F. B. (2014). Adherence to healthy lifestyle and risk of gestational diabetes mellitus: prospective cohort study. *BMJ.* 349: g5450.

Zhang, Q., He, M., Wang, J., Liu, S., Cheng, H., Cheng, Y. (2015). Predicting of disease genes for gestational diabetes mellitus based on network and functional consistency. *Eur J Obstet Gynecol Reprod Biol.* 186:91-6.

Zhao, L., Triche, E. W., Walsh, K. M., Bracken, M. B., Saftlas, A. F., Hoh, J., Dewan, A. T. (2012). Genome-wide association study identifies a maternal copy-number deletion in *PSG11* enriched among preeclampsia patients. *BMC Pregnancy and Childbirth.* 12: 61.

KASUTATUD VEEBIAADRESSID

<http://vassarstats.net/tab2x2.html>

<http://genome.ucsc.edu>

<http://www.proteinatlas.org/>

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina Anna-Maria Himma (sünnikuupäev 21.05.1993)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

„Rasedusaegsete komplikatsioonide riskiga seonduvate geenide koopiaarvu varieeruvused“,

mille juhendajad on Laura Kasak ja prof. Maris Laan,

1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates **06.02.2019** kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 26. mai 2015